2012-06

Maritime Platform Sleep and Performance Study: Evaluating the SAFTE Model for Maritime Workplace Application

Brown, Stephanie A. T.

Monterey, California. Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

MARITIME PLATFORM SLEEP AND PERFORMANCE STUDY: EVALUATING THE SAFTE MODEL FOR MARITIME WORKPLACE APPLICATION

by

Stephanie A. T. Brown

June 2012

Thesis Advisor:	Nita Lewis Shattuck
Second Reader:	Lyn R. Whitaker

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704–0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202–4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704–0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2012	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE Maritime Platform Sleep and Performance Study: Evaluating the SAFTE Model for Maritime Workplace Application			**5. FUNDING NUMBERS**
6. AUTHOR(S) Stephanie A. T. Brown			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943–5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number NPS.2011.0034-AM01-EP7-A.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

Technological advances in ship systems have enhanced the capabilities of United States Naval vessels in recent years; however, these changes come with unintended consequences. Only in recent years have we begun to study the effects of motion on the work/rest patterns of human operators in environments.

The purpose of this study was to research the performance issues related to motion in combination with the reduction of staffing onboard naval vessels. This study supports previous findings that increased motion at sea causes a decrease in sleep quality and increase in perceived fatigue. It also confirms that reaction time decreases under motion conditions.

Additionally, this study addressed concerns about the analytical approach used to assess actigraphic data and self-reported work/rest patterns in operational environments. This thesis examined the Fatigue Avoidance Scheduling Tool interface, determining that its performance predictions are dependent upon the assumptions used to score and smooth the data prior to transfer into the interface. The actual performance compared to the FAST™ performance predictions that uses the Sleep, Activity, Fatigue, and Task Effectiveness mathematical model, indicated that the model's reservoir depletion/replenishment rate did not adequately account for the effect of long-term fragmented sleep as seen in the operational maritime environment.

14. SUBJECT TERMS Motion effects on performance, Reduced manning, Performance in maritime environments, SAFTE model, FAST™ interface assessment, Sleep restrictions, Fatigue, Shiftwork			**15. NUMBER OF PAGES** 109
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540–01–280–5500 Standard Form 298 (Rev. 2–89)
Prescribed by ANSI Std. 239–18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

MARITIME PLATFORM SLEEP AND PERFORMANCE STUDY: EVALUATING THE SAFTE MODEL FOR MARITIME WORKPLACE APPLICATION

Stephanie A. T. Brown
Lieutenant, United States Navy
B.S., Cornell University, 2004

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN HUMAN SYSTEMS INTEGRATION

from the

NAVAL POSTGRADUATE SCHOOL
June 2012

Author: Stephanie A. T. Brown

Approved by: Nita Lewis Shattuck, PhD
Thesis Advisor

Lyn R. Whitaker, PhD
Second Reader

Robert F. Dell, PhD
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Technological advances in ship systems have enhanced the capabilities of United States naval vessels in recent years; however, these changes come with unintended consequences. Only in recent years have we begun to study the effects of motion on the work/rest patterns of human operators in environments.

The purpose of this study was to research the performance issues related to motion in combination with the reduction of staffing onboard naval vessels. This study supports previous findings that increased motion at sea causes a decrease in sleep quality and increase in perceived fatigue. It also confirms that reaction time decreases under motion conditions.

Additionally, this study addressed concerns about the analytical approach used to assess actigraphic data and self-reported work/rest patterns in operational environments. This thesis examined the Fatigue Avoidance Scheduling Tool interface, determining that its performance predictions are dependent upon the assumptions used to score and smooth the data prior to transfer into the interface. The actual performance compared to the FAST™ performance predictions that uses the Sleep, Activity, Fatigue, and Task Effectiveness mathematical model, indicated that the model's reservoir depletion/replenishment rate did not adequately account for the effect of long-term fragmented sleep as seen in the operational maritime environment.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION 1
A. BACKGROUND 1
B. OBJECTIVES 2
C. THESIS ORGANIZATION 2

II. LITERATURE REVIEW 5
A. FATIGUE EFFECTS ON PERFORMANCE 5
B. OPERATIONAL REQUIREMENTS ON MARITIME PLATFORMS 7
C. SLEEP AND PERFORMANCE TESTING 9
D. SLEEP, ACTIVITY, FATIGUE, AND TASK EFFECTIVENESS (SAFTE) MODEL PARAMETERS 11
E. FATIGUE AVOIDANCE SCHEDULING TOOL (FAST^{TM}) USER INTERFACE 13
F. SMART SHIP PLATFORM MANPOWER 14

III. DATA AND METHODOLOGY 17
A. OVERVIEW 17
1. Objectives 17
2. Approach 17
B. DATA COLLECTION 18
C. PARTICIPANTS 19
1. Demographics 19
2. Safety 20
3. Participant Tracking 20
D. APPARATUS 20
1. Self-Reported Questionnaire 20
2. Actigraphy 21
3. Activity Logs 22
4. Performance Tests 22
a. Switching Test 22
b. Psychomotor Vigilance Test 23
E. DATA QUALITY AND CLEANING 23
1. Self-Report Questionnaires 23
2. Actigraphy and Activity Logs 24
3. FAST^{TM} Export and Analysis Process 26
4. Performance Tests 31

IV. RESULTS AND ANALYSIS 35
A. HYPOTHESES 35
B. SLEEP DATA 36
1. Participant Descriptive Statistics 36
2. Actigraphic Data 38
3. Questionnaire Responses 40
4. Sleep Regression Analysis 45

C. PERFORMANCE DATA 48
1. Actual Performance by Sea State 48
a. Vigilance Performance 48
b. Actual Cognitive Performance 49
c. Performance Regression Models 54
2. Actual Performance Compared to Predicted Performance Model 57

V. DISCUSSION 59
A. EFFECTS OF MOTION ON SLEEP AND PERFORMANCE 59
B. DIFFERENCES BETWEEN FAST™ PREDICTIONS AND ACTUAL PERFORMANCE 60
1. Inport Versus Underway Correlations 60
2. SAFTE Model Improvements 61
C. LIMITATIONS — PROTOCOL AND SURVEY FIELDING METHODS REVIEW 62

VI. CONCLUSIONS AND RECOMMENDATIONS 65
A. EFFECTS OF MOTION ON SLEEP QUALITY 65
B. EFFECTS OF MOTION ON PERFORMANCE 65
C. ACCURACY OF THE SAFTE MODEL FOR MARITIME ENVIRONMENT APPLICATIONS 66

APPENDIX. ACTIGRAPHY DATA AND FAST™ ANALYSIS 67

LIST OF REFERENCES 83

INITIAL DISTRIBUTION LIST 89

LIST OF FIGURES

Figure 1. Schematic of the SAFTE model (From Eddy & Hursh, 2006)........................12
Figure 2. Operational schedule during data collection period..19
Figure 3. Actiwatch (From Respironics Inc., 2009)..21
Figure 4. Switching test display (From Reeves et al., 2007)..23
Figure 5. Number of test questionnaires per day...24
Figure 6. Example of actigraphy data imported from one actiwatch 25
Figure 7. Actigraphy participation collected per day 26
Figure 8. FAST^{TM} data display of pure actigraphy data import....................................27
Figure 9. FAST^{TM} data display of manually imported self-reported activity log data....27
Figure 10. FAST^{TM} output mean effectiveness (%) prediction spectrum 29
Figure 11. Quality of actigraphy data per day...30
Figure 12. Switching test mean reaction time by date 32
Figure 13. Mean reaction time over time for a) Manikin test b) Math test.......................32
Figure 14. Psychomotor Vigilance Test (PVT) by date 33
Figure 15. Rank of participants 37
Figure 16. Caffeinated beverage intake per day..37
Figure 17. Mean number of sleep episodes per day by ship status with standard deviation bars 40
Figure 18. Mean performance across trial number for Switching test and PVT (blocked on participant) 50
Figure 19. Comparison of residuals between the a) fixed model and the b) mixed-effects model 55

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

Table 1. Matched pairs t-test on FAST™ predicted effectiveness outputs with the t statistic (t), p-values (p) and sample correlations (r) 28
Table 2. Paired t-test on performance models and actual PVT results for the inport period with the t statistic (t), p-values (p), and sample correlations (r) 29
Table 3. Actigraphy determined sleep metrics by ship status (blocked on participant) 38
Table 4. Actigraphy determined sleep metrics by sea state 39
Table 5. Sleep problems responses by motion condition using Wilcoxen Rank Sum test 41
Table 6. Relationship between the mean sleep metrics and reported fatigue or sleepiness (blocked on participant) 42
Table 7. Relationship between the average amount of sleep received in previous 24 hours and fatigue, by sea state 43
Table 8. Relationship between mean sleep metrics and fatigue, by sea state (blocked on participant) 44
Table 9. Correlations between fatigue, sleepiness, and sleep received in the previous 24 hours, hours awake, and sleep reservoir, by sea state 45
Table 10. Daily sleep quantity random effects model summary with F-test statistic to test for effects..46
Table 11. Sleep quantity per sleep period random effects model summary with F-test statistic to test for effects 47
Table 12. Sleep quality random effects model summary with F-test statistic to test for effects 47
Table 13. Vigilance by ship status (two-sample t-test with block on participant)...........48
Table 14. Vigilance performance by ship motion (two-sample t-test, blocked on participant) 48
Table 15. Cognitive performance by ship status (two-sample t-test with blocked on participant) 49
Table 16. Cognitive performance by ship motion (two-sample t-test with blocked on participant) 49
Table 17. Relationship between cognitive performance and sleep reservoir level, by ship status...51
Table 18. Correlation between cognitive performance and reservoir level, by sea state 52
Table 19. Relationship between cognitive performance and sea state, by sleep reservoir level..53
Table 20. Relationship between cognitive performance and sleep received in the previous 24 hours, by sea state 53
Table 21. Vigilance performance random effects model summary with F-test statistic to test for effects...56
Table 22. Cognitive performance random effects model summary with F-test statistic to test for effects 56

Table 23. Correlation between predicted and actual vigilance performance (PVT)........57
Table 24. Correlation between predicted and actual cognitive performance (Switching MRRT overall) 58

LIST OF ACRONYMS AND ABBREVIATIONS

ANAM	Automated Neuropsychological Assessment Metrics
FAST™	Fatigue Avoidance and Scheduling Tool
FRRT	Fastest Reciprocal Reaction Time
HSI	Human Systems Integration
HSS	High Sea State
IRB	Institutional Review Board
LCS	Littoral Combat Ship
LPD	Landing Platform Dock
LSS	Low Sea State
MIF	Motion Induced Fatigue
MPT	Mathematical Processing Task
MRRT	Mean Reciprocal Reaction Time
NATO PAQ	NATO Performance Assessment Questionnaire
NPS	Naval Postgraduate School
PAWS	Performance Assessment Workstation
PSG	Polysomnography
PVT	Psychomotor Vigilance Test
SAFTE	Sleep, Activity, Fatigue, and Task Effectiveness
SRRT	Slowest Reciprocal Reaction Time
SSS	Stanford Sleepiness Survey
SWS	Slow Wave Sleep
U.S.	United States

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The Fleet Combatant Commanders have said that a 500-ship fleet is the base level required to service the demands of our force. Recently, in an effort to reduce spending in a tight fiscal climate, Defense Secretary Leon Panetta decided to reduce our fleet size to 300 vessels. This reduction in size requires our ships to be multipurpose, non-manpower intensive, and efficient. However, with efficiencies can come unintended consequences. In many cases, the vital role of human systems integration (HSI) is overlooked in the redesign of ship systems when balancing the tradeoffs between manpower, performance, and cost.

There is little known about the effects of maritime environments on performance and sleep quality in operational settings. The main objective of this study is to evaluate sleep habits and task performance of crewmembers onboard a “smart” combatant warship. Actual performance of crewmembers was compared to their predicted performance as derived from the Fatigue Avoidance and Scheduling Tool™ (FAST™) software. The sleep patterns of crewmembers were analyzed over a 14-day period, through the use of wrist worn piezoelectric accelerometers or actiwatches. The data were then used to generate estimates of cognitive effectiveness. Using the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) model in the FAST™ interface, actual performance was measured with the psychomotor vigilance test (PVT) and the Automated Neuropsychological Assessment Metrics (ANAM) Switching cognitive tests. This study compared the predicted performance based on the estimated fatigue levels using the SAFTE model with the actual performance levels in the maritime environment. It was expected that the SAFTE model would need adjustments in order to account for the higher energy expenditure and increased wake hour fatigue levels due to the maritime environment.

The performance of 21 participants was observed during inport and underway operations at multiple sea states. This study supports previous findings that increased motion on maritime platforms causes a decrease in sleep quality and increase in perceived fatigue. Activity counts during sleep periods increased significantly, indicating sleep fragmentation. It also confirms that reaction time as measured by the PVT decreases during at sea periods.

Additionally, this study addressed concerns about the analytical approach used to assess actigraphic data and self-reported work/rest patterns in operational environments. This thesis systematically examined the FAST^{TM}, determining that its performance predictions are directly dependent upon the subjective assumptions used to score and smooth the data prior to transfer into the FAST^{TM} interface. The actual performance results compared to the FAST^{TM} performance predictions that use the SAFTE mathematical model indicated that the model's reservoir depletion or replenishment rate did not adequately account for the effect of long-term fragmented sleep as seen in the operational maritime environment.

Further studies should be conducted to investigate the effects of motion on performance in order to create an accurate fatigue and manpower model for fleet-wide shipboard implementation. These studies should account and control for sleeping environmental factors such as noise, temperature, and humidity, as well as other factors including increased sleep duration during high sea states, use of seasickness medication, and caffeine use may have affected the performance results. Finally, future studies should have a baseline study in ideal sleeping conditions for comparison to ensure the participants are not suffering from total or partial chronic sleep deprivation, which may change their sleep cycle.

ACKNOWLEDGMENTS

I would first like to express my gratitude to all those who helped me complete this thesis research. I would like to thank my advisors, Dr. Nita Shattuck and Dr. Lyn Whitaker, for their time and guidance along the way. Dr. Shattuck's continuous enthusiasm and energy inspired me throughout my time in the Human System's Integration program. I would also like to thank NPS President Daniel Oliver, for his constant support of higher education and providing me with the opportunity to accomplish my aspirations while serving as his Flag Aide. Additionally, I would like to thank the testing team members from Naval Surface Warfare Center Panama City, and Johns Hopkins University for collecting the data used in this thesis, as well as the crewmembers who participated in the study. I would like to also thank my family for their years of support and encouragement. And finally, thank you to my loving husband, Chris, for his invaluable friendship, positive outlook, and endearing sense of humor.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

"The military will be smaller and leaner, but it will be agile, flexible, ready and technologically advanced; it will be cutting edge," Defense Secretary Leon Panetta told reporters at the Pentagon as he unveiled details about the fiscal 2013 budget (Brannen, 2012). The Fleet Combatant Commanders have said that a 500-ship fleet is the base level required to service the demands of our force. Recently, in an effort to reduce spending in a tight fiscal climate, Panetta decided to reduce our fleet size to 300 vessels. This reduction in size requires our ships to be multipurpose, non-manpower intensive, and efficient. However, with efficiencies can come unintended consequences. In many cases, the vital role of human systems integration (HSI) is overlooked in the redesign of ship systems when balancing the tradeoffs between manpower, performance, and cost.

The country's naval forces have evolved over the years, both in mission and capability. The original warship relied solely on the external environmental factors and onboard manpower to be an effective force. The speed and accuracy of the ship was entirely dependent on the talent of its captain and the readiness of its crew. In today's world of technological advancements, ship operations are not entirely reliant upon brute force. Ship systems and missions have evolved dramatically over the years, changing the manpower and personnel needs. Most recently, increased levels of automation have led to major changes in many work environments onboard ships, having a profound effect on working patterns and job performance (Woods, Sarter, & Billings,1997). This increase in automation has created some major benefits for the overall human-system performance, however it has also introduced a new class of human performance concerns onboard maritime platforms (Sarter & Woods, 1995). Reduced manpower puts additional requirements on the remaining crew, increasing necessary skills, training, and time. This cross-rating leads to potentially disruptive work cycles and reduced quantity and quality sleep in the crew. The combination of reduced manning and sleep debt accumulation presents a serious challenge to new modular mission platforms.

B. OBJECTIVES

This thesis explores the application of the reduced manning model associated with the new modular maritime platforms. By comparing actual crew performance levels with the predicted performance based on sleep obtained as measure through actigraphy, I analyze the accuracy of current fatigue models used for manning maritime platforms. The study evaluates the models to see if they appropriately account for the effects of motion on overall crew sleep efficiency and performance effectiveness. This thesis discusses the following objective areas:

Sleep Disturbances and Physical Activity: Investigate the relationship between sleep loss and physical activity and the effects of rough sea conditions on physical activity and sleep disturbances.

Fatigue: Investigate the extent to which fatigue is related to sleep time and sleep quality.

Performance: Quantify cognitive performance impacts from fatigue induced by ship motion.

Validation of Previous Data: Compare previous Fatigue Avoidance Scheduling Tool™ (FAST™) model predictions based on actigraphy with actual participant performance.

SAFTE Model Improvements: Based on results of the validation, make recommendations to improve both the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) model and its application through the FAST™ interface.

C. THESIS ORGANIZATION

Chapter I describes the background of the U.S. Navy's reduced manning concept and the drivers motivating the development and implementation of new modular combatant maritime platforms. Chapter II contains a literature review of sleep and fatigue in civilian and military settings, shipboard operational factors, performance testing in sleep studies, the evolution, validation, and application of the SAFTE model and FAST™ user interface. Chapter III describes the methodology and data collection techniques used to quantify sleep quality and performance. The results and implications of reduced quality sleep on cognitive performance in maritime environment are discussed in Chapter IV and V. Finally, the conclusions and recommendations are defined in Chapter VI.

THIS PAGE INTENTIONALLY LEFT BLANK

II. LITERATURE REVIEW

Sleep is a vital process for the survival of all organisms. It is a time where the body is able to heal, reorganize, and regroup. It has been shown that total sleep deprivation will lead to major health problems and will eventually be fatal (Coren, 1996; Naitoh, Kelly, & Englund, 1990). This section discusses the effects of fatigue and sleep loss on reaction time and cognitive performance. It also discusses the increase demands on the human body during maritime operations. Sleep and fatigue models have been developed and adjusted in the most recent years to help predict the effects of sleep loss on performance. This chapter reviews some of these models and their application to combat operations and maritime environments. Finally, this chapter reviews the evolution of manning requirements onboard combatant warships, and the tools used to determine performance based on manpower and sleep requirements.

A. FATIGUE EFFECTS ON PERFORMANCE

It has been shown in many studies that rotating watch assignments can have serious effects on an individual's circadian rhythms, particularly their sleep cycle (Colquhoun & Folkard, 1985; Goh et al., 2000; Rutenfranz et al., 1988; Sack et al., 2007). Even relatively moderate sleep restrictions can change sleep architecture and seriously impair waking neurobehavioral functions in healthy adults (Van Dongen et al., 2003). Currently, Naval vessels use continuous five-hour bridge watches utilizing a rotating watch schedule. The number of watch sections is directly dependent on the number of qualified officers and the command leadership organizational culture onboard. It is typical to have a schedule of four sections, however this number is decreasing due to the reduction in manpower onboard surface ships. Studies have shown that a rotating watch system causes an increase in fragmented sleep in individuals, leading to a degree of rhythm disintegration during prolonged periods (Colquhoun & Folkard, 1985; Hakola & Härmä, 2001). This reduction in quality sleep increases overall fatigue, affecting the sailor's performance levels, particularly on simple, repetitive and long-duration tasks where motivation levels are the lowest (Horne, 1985; Rutenfranz et al., 1988). In addition, the stress of sleep loss can cause asthenopia (eyestrain), a decrease in contrast sensitivity, and an overall decrease in visual efficiency (Quant, 1992), as well as other major health concerns (Naitoh et al., 1990). Since a typical watch consists of the task of continuous monitoring of visual display terminals, a decrease in visual efficiency may lead to misinterpretations of the displays increasing the risk of accidents and dangerous situations.

Time of day has great impacts on performance as well. Night shifts cause increased mental fatigue and confusion, as well as decreased arousal and activity levels (Luna et al., 1997). Other studies show that performance levels often follow the individual's core temperature level (Folkard & Monk, 1985; Sacks et al., 2007), thus decreasing performance during night shifts when the core body temperature is at its lowest. Colquhoun and Folkard (1985) tested vigilance in individuals, finding a significant increase in reaction times when body temperature rose, and a drastic decrease in performance between 0000–0400 when body temperature is at its lowest. Daytime circadian rhythms cause a rise in performance, counteracting sleep deprivation effects. However, at night the rhythms fall, thus adding to the effect of sleep deprivation on performance with the lowest point at 0400 (Folkard & Monk, 1985; Akerstedt, 1990). Individuals never can truly adjust their sleep patterns when subjected to a rotating work-rest schedule, therefore always being moderately sleep deprived, exhibiting severe sleepiness and a reduced performance capacities (Akerstedt, 1990; Colquhoun & Folkard, 1985). Extended duty shifts (10–12 h) have become increasingly popular because they maximize the time off from work (Sack et al., 2007) but may have unforeseen performance consequences. Shift work creates potentially hazardous implications on the safety of Naval vessels and their crews.

Another aspect that has effects on performance is an individual's arousal level. There are several ways to counter poor sleep quality and reduced quantity of sleep. Caffeine ingestion has been shown to increase visual vigilance, reaction time and alertness, while reducing the effects of sleep deprivation by increasing mental arousal (Lieberman et al., 2002). Energy levels affect arousal as well. The ingestion of food decreases the core body temperature, thus decreasing arousal levels and reducing performance (Folkard & Monk, 1985). This affect may partially account for the "postlunch dip" (a reduction in energy levels in the afternoon, directly following midday meal). However, food consumption will eventually increase energy levels, improving performance. Planned napping is another form of intervention to counter the effects of poor quality sleep resulting from shift work. In a study by Sallinen et al. (1998), it was shown that napping resulted in improved reaction times in the early morning periods, increasing alertness. Other studies have shown that napping prior to night shifts, especially in combination with caffeine, have resulted in reduced accidents and improved alertness as assessed by psychomotor vigilance testing (Purnell et al., 2002; Schweitzer et al., 2006).

While onboard ships, sailors expend more energy simply from continuously balancing and adjusting to the ship's degrees of pitch and roll. This extra expended energy called motion induced fatigue (MIF), can increase physical fatigue in an individual nearly twice as much as working in a stable environment causing reductions in performance (Heus et al., 1988; Wertheim, 1998). This extra energy expenditure would also change the rate of sleep reservoir depletion. Most sleep-performance models describe the homeostatic process as a simple reservoir in which performance capacity increases exponentially during sleep and decays linearly during wake periods (Akerstedt & Folkard, 1997; Johnson, 2004). These notional processes have been successful at describing the relationship between sleep and performance under conditions of irregular sleep patterns, jet lag, and short periods of complete sleep deprivation (Akerstedt & Folkard, 1997; Achermann & Borbély, 2003; Balkin et al., 2000). However, none of the studies have looked at the homeostatic process in a maritime environment with irregular motion and energy expenditures.

B. OPERATIONAL REQUIREMENTS ON MARITIME PLATFORMS

Maritime platforms provide a unique environment for our workforce. Operational requirements at sea are strained by the additional stress of the environmental factors. Human error becomes the dominant cause for approximately 75–96% of the accidents in a maritime setting (Rothblum, 2002). Fully understanding the additional environmental stressors of shipboard environments could help improve finding the root cause of the maritime casualties, ultimately reducing the probability of human error.

There are assumptions that working onboard a ship is more strenuous than comparable work ashore. It has been shown that during pitch and roll movements of a ship motion simulator platform, the energy expenditure for a walking task increased by 30% as compared to stationary control conditions (Heus et al., 1998; Wertheim, 1998). This finding implies a greater muscular effort and workload when performing tasks onboard maritime platforms, resulting in increased physical fatigue. Complex tasks can be greatly affected by shipboard motion. Several studies using simulated shipboard environments have been conducted previously, looking at complex tasks typical of real naval operations requiring decision making based on radar image interpretation and memorization involving both cognitive skills, perceptual skills, and fine motor coordination skills. These studies showed that in a moving environment, there is a small but significant reduction in information transfer during operations (Heldsdingen, 1996; Wertheim & Kistemaker, 1997). Other shipboard studies of cognitive tasks have had inconclusive results, indicating that cognitive skills may not be directly affected by ship movements, but may have some indirect effects (Crossland & Lloyd, 1993; Gillard & Wientjes, 1994; Wertheim, 1995). Most of these studies were short-term, and did not explore the effects of different levels of shipboard movement, or sea states.

Sleeping environments are extremely important in order to ensure high efficiency during rest periods. In an operational environment, it is especially critical to ensure good sleep quality for the whole crew since decision-making relies on inputs from the entire watchteam and the results could be catastrophic if a single interpretation is incorrect. Shipboard design of sleeping quarters has improved over the years to increase user habitability (Meere & Grieco, 1996). For example, the design of the new San Antonio class Landing Platform Dock (LPD) ship improved sleeping quarters, providing better ventilation, storage space, and increased head clearance (Defense Industry Daily, 2012). Another example is the new Littoral Combat Ship (LCS), which provides staterooms for all crew, not just officers. This design feature increases privacy and comfort, while decreasing noise, smells, and other environmental factors that can cause sleep disturbances. As seen in previous studies, these habitability factors can greatly affect the sleep quality received by an individual.

However, no habitability design improvements can fully counter the effects of motion on sleep. Motion and vibrations have been shown to increase the number of sleep bouts and wake periods during the night, preventing the human body from reaching deeper stages of sleep required for reservoir restoration (Calhoun, 2006). In addition, sleep stages have been shown to change in length in a maritime environment. Specifically, the time spent in sleep stage one significantly decreases under motion conditions, and the overall sleep cycle tends to lengthen (O'Hanlon et al., 1977). These changes in the sleep architecture can have a major effect on the rejuvenating nature of sleep, thus decreasing performance. New hull designs should closely consider these factors when under review.

C. SLEEP AND PERFORMANCE TESTING

Sleep and performance testing are both difficult areas of objective evaluation due to confounding external factors. Both fields have grown in recent years. This study utilizes sleep and performance testing methods based on the recommendations from previous studies.

The measurement of sleep is an area of study that has much uncertainty. Current popular objective methods of recording sleep include polysomnography (PSG) and wrist worn actigraphy. Recording brain wave activity through PSG provides more insight to the neurobehavioral functions and sleep physiology than actigraphy. It is appropriate for laboratory settings, but essentially impractical during field-testing. Actigraphy is a simpler method, which measures the movement of the participant during sleep. Combining actigraphy with our knowledge about sleep cycles, we can make fairly accurate conclusions about an individual's sleep quantity and quality.

Previous studies have questioned the subjectivity of sleep scoring programs and actigraphic algorithms (Cole et al., 1992; Jean-Lousie et al., 2001). Kripke et al. (2010) conducted a study on the performance the Actiwatch product commercially owned by Respironics, now Philips Electronics. The study showed that the Actiware automatic scoring was not very accurate in comparison to PSG readings.

Performance testing is also difficult due to questions of test validity and subjectivity. Most testing requires a baseline at peak performance in order to conduct a within subjects analysis. A baseline study requires strict controls in an experimental design, which are difficult to obtain in an operational setting. Although a controlled laboratory setting would help isolate effects, the operational setting is valuable because it allows us to obtain additional interactions and environmental factors that may not have been foreseen otherwise.

The psychomotor vigilance test (PVT) is among the most widely used measures of alertness. The PVT is used due to its high sensitivity to sleep deprivation. The advantage of using the PVT over other cognitive batteries is its simple way to track behavioral alertness changes caused by inadequate sleep, without the confounding effects of aptitude and learning (Basner & Dinges, 2011; Graw et al., 2001). However, there are differing views on its ability to predict cognitive performance levels. Some believe that alertness alone is not an adequate measure of performance due to its isolated transferability to real world performance tasks. There are many aspects of cognitive performance which are based on task type and mental process requirements. Alertness is simply a measure of reaction time. Complex cognitive performance requires mental processing. Sleep deprivation and fatigue affect parts of the brain differently.

The Automated Neuropsychological Assessment Metrics (ANAM) is a series of computerized tests and test batteries designed by the U.S. Military for testing cognitive processing in a variety of contexts that include neuropsychology, fitness for duty, neurotoxicology, pharmacology, military operational medicine, human factors engineering, aerospace and undersea medicine and sports medicine (Reeves et al., 2007). The Switching Task, part of the ANAM, is a combination of a spatial orientation test (manikin test) and a computational reasoning test. The test is designed to be an executive function task that requires the ability for mental flexibility and shifting set (Reeves et al., 2007). It was derived from the Performance Assessment Workstation (PAWS) battery, which measured short-term memory, spatial processing, attention, tracking, and task timesharing. Eddy et al. (1998) used the PAWS battery in a space setting to test the effects of microgravity environments on cognitive performance.

The Manikin Task is the spatial orientation task requiring the participant to determine the location of an object being held by an image of a man. The man's orientation is varied throughout the test. The Mathematical Processing Task (MPT) requires the participant to do basic arithmetic presented on the display, and determine if the answer is greater or less than five. Performance degradation has previously been seen in mathematical processing tasks when participants are fatigued (Harvile et al., 2007).

D. SLEEP, ACTIVITY, FATIGUE, AND TASK EFFECTIVENESS (SAFTE) MODEL PARAMETERS

Fatigue has been an area of interest in the study of work over the past three decades. In the age of globalization with advancements in technology and need for military preparedness, optimal human performance is required 24/7. Many models have been developed and tested to predict human performance based on fatigue levels as determined by sleep achieved and circadian desynchronization. These models try to quantify the influence and impact of these factors on sleep propensity, wake alertness, and overall performance. Since the early '90s, many studies have been conducted to model the fatigue and performance relationship. Many of these models have been conducted in land-based environments such as army operations (Hursh et al., 2004). Mallis et al. (2004) did a comparison study of the key features of seven biomathematical models currently in development or in commercial use, concluding that the SAFTE model best encompassed the factors leading to fatigue and most accurately predicted performance.

The SAFTE model is designed to determine how the time of day or circadian rhythms, and sleep-wake patterns influence cognitive capacity and risk of performance error (Hursh, 2003). It was developed for both military and commercial use, and has been applied by schedulers and managers in the U.S. Air Force and the Federal Railroad Administration to optimize operations through the Fatigue Avoidance Scheduling Tool (FAST™) (Eddy & Hursh, 2006; Mallis et al., 2004). As seen in the schematic shown in Figure 1, the model encompasses three processes: the homeostatic process, sleep inertia, and the circadian process (Hursh et al., 2004). The combined model for performance effectiveness as expressed as a percent of baseline is given by Eq. 1:

$$E_t = 100 * (R_t/R_C) + C_t + I \qquad \text{Eq. 1}$$

where $100 * (R_t/R_C)$ is the reservoir level; C_t is the circadian process; and I is the transient inertia. The circadian process component accounts for the variability of effectiveness with increased sleep debt.

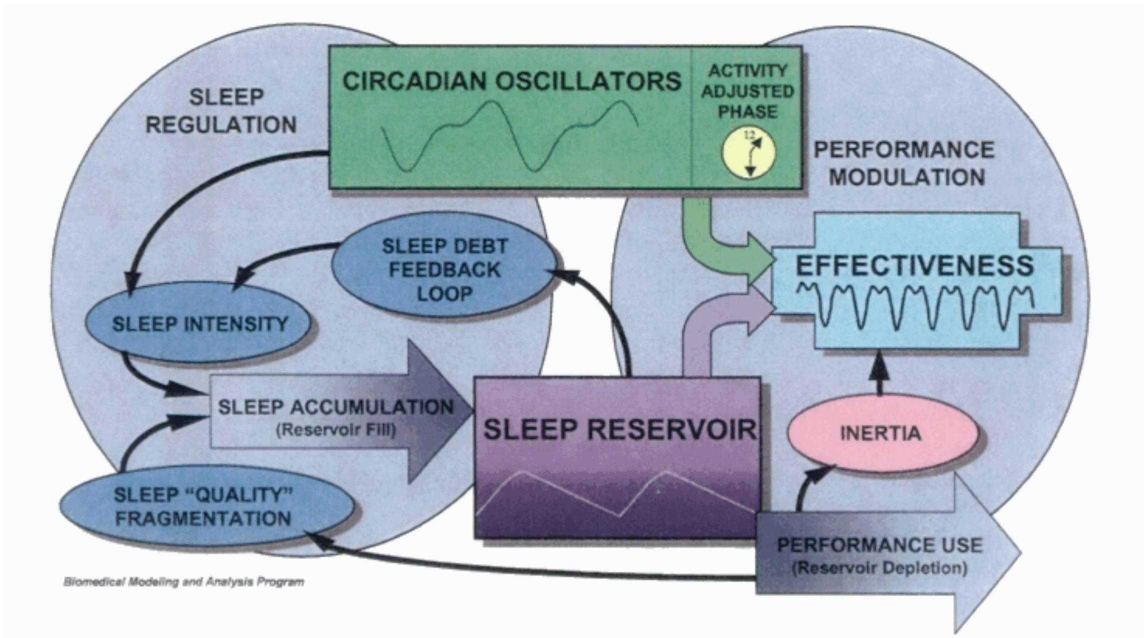


Figure 1. Schematic of the SAFTE model (From Eddy & Hursh, 2006)

As with most biomathmatical models, the SAFTE model has some limitations. First, the model was based on a college-aged student population during laboratory settings (Mallis et al., 2004). The model's accuracy may be limited when applied to other population types, such as operational military members.

Another limitation is in the area of the algorithm's predicted performance output. The model makes performance predictions based on vigilance performance metrics, which test reaction time. However, ultimately all models are judged based on the usefulness of their predictions of performance in areas of greatest interest to the user (Hursh et al., 2004). Hursh and others used PVT measures to validate and adjust the SAFTE model based on a general understanding that this is the "golden measure" for cognitive performance (Hursh et al., 2004). Eddy and Hursh (2006) have optimized the

SAFTE model for PVT speed because they claim the PVT showed greater sensitivity and degradation across all restricted sleep conditions than the combined cognitive tests.

Another potential issue with the SAFTE model is its components and application. The SAFTE model predicts the average results with precision, but its accuracy has not been tested in an operational setting with unusually high sleep-wake cycles such as those experienced by shipboard sailors. Eddy and Hursh (2006) recognized that the model's error in effectiveness prediction is large at extreme and chronic sleep restriction situations due to a considerable amount of variance in performance between individuals (up to 60% difference between participants). It also does not include the effects of physical work, workload, or level of interest in task (Mallis et al., 2004). In addition, it does not have the capacity to change the reservoir level depletion rate, which is predicted to be higher onboard maritime platforms due to the increased energy expenditure rate to compensate for the platform motion as found by Heus, Wertheim and Havenith (1998).

E. FATIGUE AVOIDANCE SCHEDULING TOOL (FAST™) USER INTERFACE

The Fatigue Avoidance Scheduling Tool (FAST™) software, using SAFTE model, has recently been adjusted to account for most recent sleep history in the projected population. The model integrates information about circadian rhythms, cognitive performance recovery rates associated with sleep and decay rates associated with wakefulness, and cognitive performance effects associated with sleep inertia (Eddy & Hursh, 2006).

However, the interface has been enhanced to allow SAFTE to predict group variances around the mean (Eddy & Hursh, 2006). This means that the FAST™ interface can effectively predict performance for a group, but is less accurate for individual predictions. Individual performance varies greatly based on a plethora of factors specific to each person to include the differences in sensitivity to sleep deprivation. Because of this enhancement, the FAST™ interface has a large error rate for individual predictions at the extreme ends of the sleep spectrum (i.e., individuals experiencing greater sleep restrictions and chronic sleep restrictions). Eddy and Hursh (2006) stated that in the Sleep Dose Response study, the mean speed on the PVT for individual subjects ranged from

80% to 20% of baseline. This large variability limits the application of this model and interface to predict individual readiness, especially in extreme environments and operational conditions such as onboard maritime platforms.

The FAST™ interface has evolved through many versions in order to enhance its capabilities and usability. Over its many improvement iterations, it has added features to allow users to adjust sleep environments, display shifts in time zones, and analyze the effects of sleep aides on performance. However, the model and interface have not accounted for differences in work environments. For example, it may not be an accurate predictor for maritime environments because it does not account for the greater amount of energy expenditure during wake hours when working in a dynamic environment. It has been shown in a previous study that the muscular effort needed for maintaining balance when walking or working on a pitching/rolling platform results in a significantly higher workload than similar work on a stable floor (Heus et al., 1998). This increased fatigue observed when a task is performed on a moving platform should be accounted for in the model by a larger reduction in performance effectiveness during wake hours. This additional workload will vary between ship platform based on the severity of motion and stability of the hull.

F. SMART SHIP PLATFORM MANPOWER

"Ships of the next century will have automation, smarter systems and fewer sailors," (Meere & Grieco, 1996). The Navy has moved towards design of "smart" ships, capable of operating with fewer personnel onboard. The increase in automated, self-sustaining technology has allowed for a reduction in manpower due to the reduced number of tasks and corrective maintenance associated with the systems. The new modular mission concept for the Littoral Combat Ships (LCS) is also reducing the requirement for large core crew sizes. The LCS only requires 40 core crewmembers to operate, plus an additional 20 for the aviation detachment and 15 for the specific mission modules (U.S. Library of Congress, 2011). The limited number of mission modules reduces the number of possible missions the crew could be called to execute at any given time. This manning level reduces the capability of the ship at any given time, making it unable to perform multiple missions at once, but also reduces the required core crew size.

This manning reduction creates a more efficient ship, but also requires planning and forward deployment of the vessel and modules. It could be argued that the manpower cost savings provided by the reduced crew size does not outweigh the cost of increased forward deployed base infrastructure.

Rotating crews is the new design concept in shipboard manpower. The concept has been implemented first onboard the ballistic missile submarines with blue and gold crews. While the ship is underway for three months with one crew, the other crew is inport training. Similarly, the new LCS is utilizing the rotating crew concept. For each ship, three crews would be maintained. The LCS is using a "3–2–1" plan, where the ship is deployed for 16 months at a time, and the crews are rotated on and off at four month intervals (U.S. Library of Congress, 2011). This concept allows the Navy to utilize their assets at higher rates, while training the crews on land. However, this concept has potential for failure. Ships tend to deteriorate quickly in operational settings. With reduced numbers in crew size, the ongoing preventative maintenance that is required for upkeep may not occur at an ideal level. This could lead to the physical deterioration of the ship, as well as increased hazards for crewmembers onboard. The operational crew will likely have to conduct maintenance that was not planned for in the original ship design concept and manpower document. This will increase the workload for the small crew, and decrease their overall effectiveness.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DATA AND METHODOLOGY

A. OVERVIEW

1. Objectives

Little is known about the effects of maritime environments on performance and sleep quality in operational settings. The main objective of this study was to evaluate sleep habits and task performance of crewmembers onboard a "smart" combatant warship. Actual performance of crewmembers was compared to their predicted performance as derived from the FAST^{TM} software. The sleep patterns of crewmembers were analyzed through the use of wrist worn piezoelectric accelerometers or actiwatches. These devices were used to measure the gross motor activity of the crewmembers, which provided an indicator of sleep length and quality. The data were then used to generate estimates of cognitive effectiveness. Using the SAFTE model in the FAST^{TM} interface, actual performance was measured through psychomotor vigilance and cognitive tests. This study compared the predicted performance based on the estimated fatigue levels using the SAFTE model with the actual performance levels in the maritime environment. It was expected that the SAFTE model would need adjustments in order to account for the higher energy expenditure and increased wake hour fatigue levels due to the maritime environment.

2. Approach

This study will evaluate a number of areas using tools and methods based on the recommendations found in previous studies. The literature review discussed a number of accuracy issues and validation analysis of various methodologies when conducting sleep and performance studies. In the area of actigraphy analysis, due to the issue of poor accuracy in Actiware automatic scoring in comparison to PSG readings found by Kripke et al. (2004), this study manually scored the actigraphy data. Additionally, this study used the medium threshold level for activity counts that were considered "awake" (40 counts per minute) due to its highest level of agreement with the PSG readings (Kripke et al., 2010).

When choosing performance measures, both PVT and Switching tests were utilized. As previously discussed, PVT measures reaction time rather than cognitive performance that is required for complex tasks. Cognitive tasks, such as math computation and spatial orientation, require cognitive processes that differ from vigilance. In this study, we use both PVT and the ANAM Switching tests to validate the SAFTE model's predictive accuracy. If the Switching test has been proven robust enough to determine cognitive performance in an extreme environment with combined stressors such as space as seen in Eddy and Schifflet's study (1998), it should be reliable for shipboard use.

Finally, this study evaluated the properties of the SAFTE model to determine their accuracy in predicting performance in personnel with extreme sleep levels, observed during actual sea-based operations. This study also looked at the FAST™ interface under a new lens. The interface has been validated in land-based studies, but never for maritime application. Additionally, this study looked at the model and interface accuracy in predicting complex cognitive processing beyond vigilance.

B. DATA COLLECTION

All data were collected during a single 14-day period onboard a U.S. Naval Combatant. The operational schedule as seen in Figure 2 allowed for testing underway during low sea state conditions from 7 March, 1200 hours, to 9 March, 0900 hours. The ship was inport from 9 March, 0900 hours, until 16 March, 1600 hours. This inport period served as the baseline for the study. The testing at sea was then continued from 16 March, 1600 hours, until 21 March, 1300 hours.

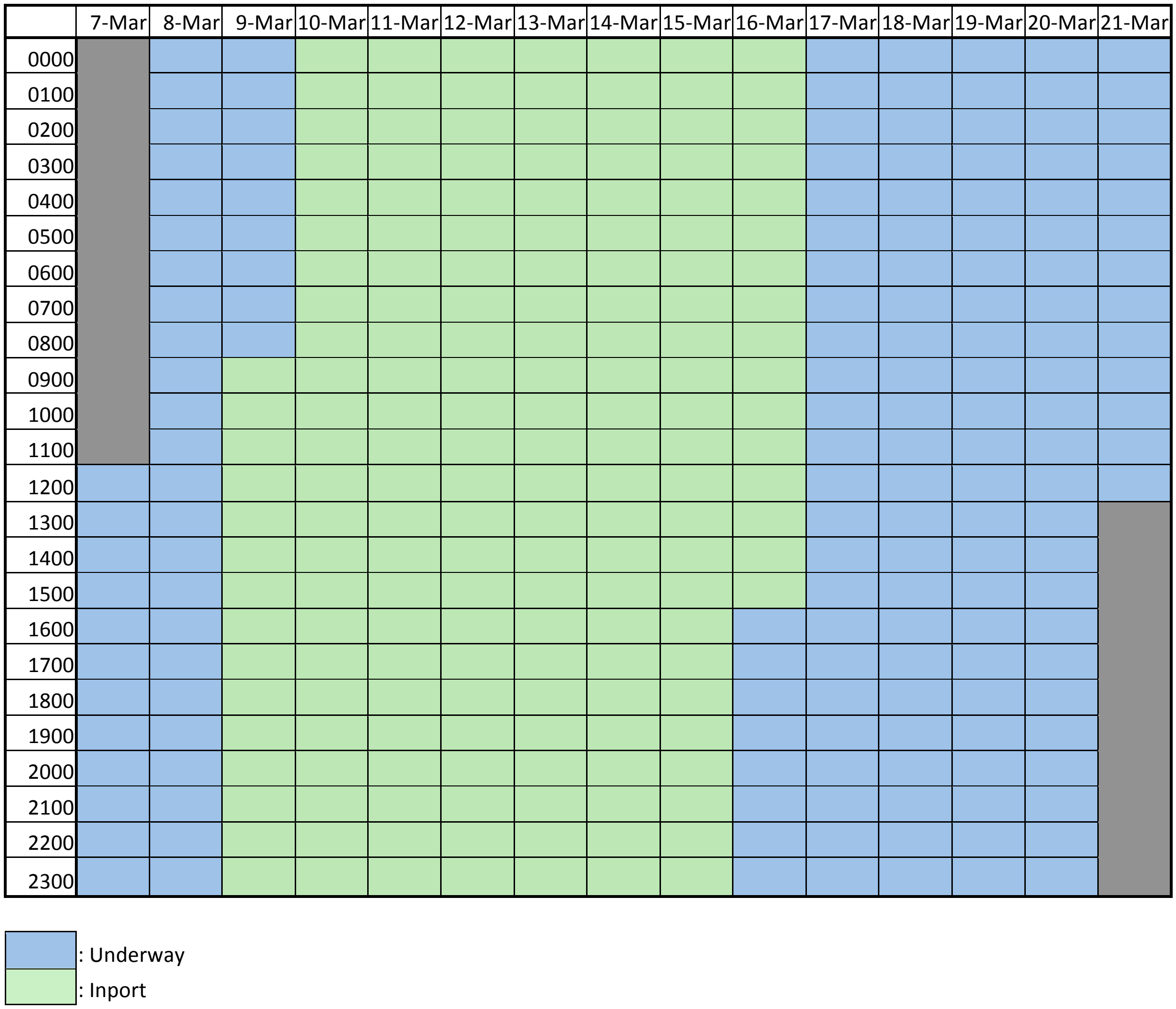


Figure 2. Operational schedule during data collection period

Crewmembers volunteered to participate in a series of data collection opportunities to include: actigraphy sleep collection, activity logs, performance tests to include Switching and PVT, and fatigue questionnaires to include NATO Performance Assessment Questionnaire (PAQ) and Stanford Sleepiness Survey (SSS).

C. PARTICIPANTS

1. Demographics

Of the 40-person crew, only 21 volunteered for the study (19 males, 2 females). The participants ranged in pay grade from E-5 to O-4. Due to random failures in the actiwatch devices as well as improper participant use, data from only 15 personnel were usable for analysis (14 males and 1 female). Section E discusses the data quality and cleaning process, and the resulting number of data sets analyzed in this study.

2. Safety

The data collection team briefed each participant on the expectations of the study. Each participant signed a voluntary participant consent, privacy act, and personal custody form. Each participant had the opportunity throughout the study to ask questions and discuss any concerns with the procedures or equipment issued. The participants were able to withdraw from the study at any time.

The NPS Institutional Review Board (IRB) approved the use of actiwatches and computer-based testing used throughout this study. Participation from the crew was not mandatory. There were no known psychological or emotional risks associated with participation in the data-collection process. The participants were exposed to the same inherent risks as onboard any U.S. Navy vessel.

3. Participant Tracking

Each participant was tracked using a randomly assigned identification code. The actiwatches and data collection devices issued to the participant each had their own device number. The list of names and corresponding identification numbers were secured by the data collection team's test director to ensure the participants' anonymity in the analysis process.

D. APPARATUS

A variety of objective measures of human performance were collected onboard the operational maritime platform in an attempt to quantify the impact of ship motion on sleep and human performance. These measures included actigraphy and performance tests. Subjective measures were also collected through the use of self-report surveys.

1. Self-Reported Questionnaire

The participants took an initial survey which collected baseline information on their work and sleep habits. Self-reported sleep information was collected throughout the testing period through the use of the NATO PAQ and SSS. These tools were designed to determine self-reported fatigue levels experienced over the previous 24 hours. In the questions, participants were asked to rate on a four-point Likert scale whether they disagree (“0”) or agree (“3”) to the statements that the “the quality of sleep was poor,” and that the “amount of time sleeping was short.” Based on these ratings, we developed the aggregate metric “Sleep problems total score,” calculated as the mean of all symptoms’ ratings. Therefore, total score ranges from zero (no symptoms experienced) to three (high level of symptoms experienced). In the last three statements of this group of questions, participants were asked to rate on a four-point Likert scale the extent that their sleep problems were caused by ship motion, seasickness, or other factors (zero = no association, three = extreme association). All questionnaires were self-administered on iPod Touch devices twice a day.

2. Actigraphy

Actigraphy was collected through the use of individual actiwatches. Actiwatches are wrist-worn piezoelectric accelerometers that collect information on the wearer’s motion (Respironics Inc., 2009). Figure 3 shows the actiwatch apparatus worn by the participants in this study.

Figure 3. Actiwatch (From Respironics Inc., 2009)

Participants were instructed to wear the watches like a wristwatch on their non-dominant arm during the entire collection period. The device documented their resting and active times throughout the testing period. The participants were allowed to remove the actiwatch for short periods of time when it interfered with their personal safety or with daily tasks. For example, some participants chose to remove their watches during cleaning duties and while in the shower. These removal periods were documented in the participants' daily activity logs and then excluded from the analysis.

3. Activity Logs

Participants were asked to maintain a log of their daily activities. Each participant had an iPod Touch that allowed them to record all daily tasks in categories such as maintenance, watch, administration, sleep, eating, and training. The logs were updated with start and stop times for the preselected activities. The purpose of this log was to allow for later comparison with the actigraphy when cleaning the actigraphy data extracted from the actiwatches.

4. Performance Tests

Performance was measured using a portion of the cognitive Automated Neuropsychological Assessment Metrics (ANAM) test battery called the Switching Test and a psychomotor vigilance test (PVT). Both tests were administered on a computer twice a day.

a. Switching Test

The ANAM Switching Test is designed to evaluate high-level decision making as well as three-dimensional spatial rotation, basic computation skills, concentration and working memory. It is designed to be an executive function task that requires the ability for mental flexibility and shifting set (Reeves et al., 2007). It was originally designed for within-subjects comparisons and does not have traditional normative group data (Reeves et al., 2006). The manikin task, located on the left side of the computer screen, is a spatial orientation task requiring the participant to determine the location of an object being held by an image of a man. The man's orientation is varied throughout the test. The mathematical processing task (MPT) on the right side of the screen requires the participant to do basic arithmetic presented on the display, and determine if the answer is greater or less than five. Between these two images, a red arrow randomly points to one side of the screen or the other (see Figure 4). The participant must complete the task indicated by the arrow. If the arrow points to the math problem, the participant must calculate the solution. If the arrow points to the manikin holding the ball and box, the crewmember is asked to determine which hand contains the object of interest. Throughout the trials, the manikin randomly rotates orientation and the objects move between the hands. Performance degradation has previously been seen in mathematical processing tasks when participants are fatigued (Harvile et al., 2007).

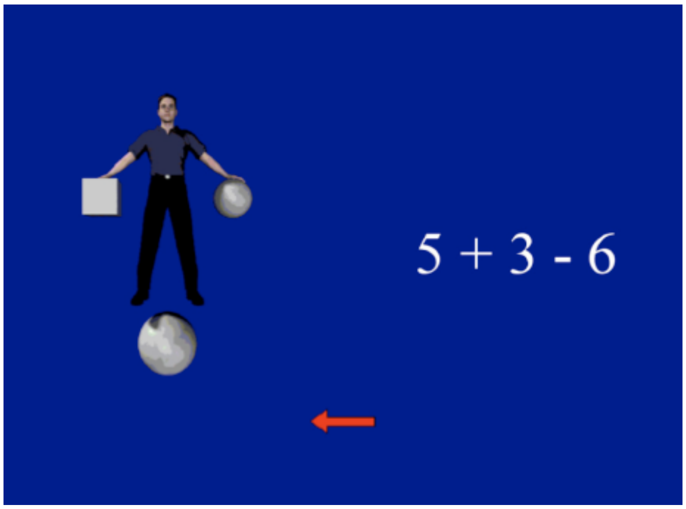


Figure 4. Switching test display (From Reeves et al., 2007)

b. Psychomotor Vigilance Test

The PVT is designed to measure reaction time of the participant (Pulsar Informatics, 2012). A blank black screen with the outline of a red rectangle is displayed to the crewmember. When numbers appeared inside the rectangle, the participant immediately hits the space bar on the computer. This test is designed for a three-minute interval test session.

E. DATA QUALITY AND CLEANING

1. Self-Report Questionnaires

The self-reported questionnaires were derived from the NATO PAQ and the SSS. These questionnaires were cleaned to remove any redundant or conflicting entries. Questionnaires of participants who did not have corresponding quality actigraphy data were not used for analysis. Analysis was based on 115 completed test questionnaires, 45 completed while inport and 70 underway. On a daily basis, the number of test questionnaires ranged from 3 to 14 (Mean=8.2, StdDev=3.49). Figure 5 depicts the distribution of questionnaire analyzed per day.

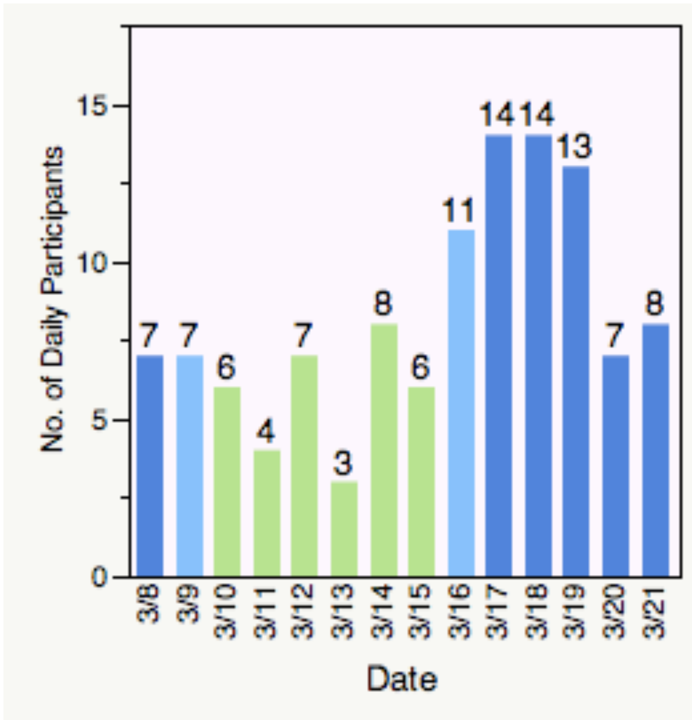


Figure 5. Number of test questionnaires per day

2. Actigraphy and Activity Logs

The actigraphy data were downloaded and evaluated for completeness from the actiwatches using Respironics Actiware version five software. First, the actigraphy database start date and time had to be adjusted to match the individual participant recorded watch return time and the common period of major activity during return transport on March 24 from 0845 to 1050 hours.

Next, the activity logs were cleaned to ensure no overlapping or conflicting data were recorded. During the beginning of the testing phase, the iPod devices had technical issues dealing with time zone alignment. The test administrators reprogrammed every iPod device, and the resulting activity logs were adjusted to reflect the proper time zone. The tasks recorded in the participant's individual activity logs were then transferred into Actiware as "Forced Wake" intervals, or "Rest" intervals. The periods of time when the actiwatch was removed were recorded as "Exclusion" intervals. In addition, "Exclusion" intervals were included for periods of high sea states as shown in the ship log with zero activity in the database, indicating that the watch was removed without logging the event. When the activity log did not show any activity, and the actigraphy appeared to have minimal activity, the interval was also recorded as "Rest." Figure 6 shows an example of actigraphy data from one participant (all actigraphy data in the Appendix).

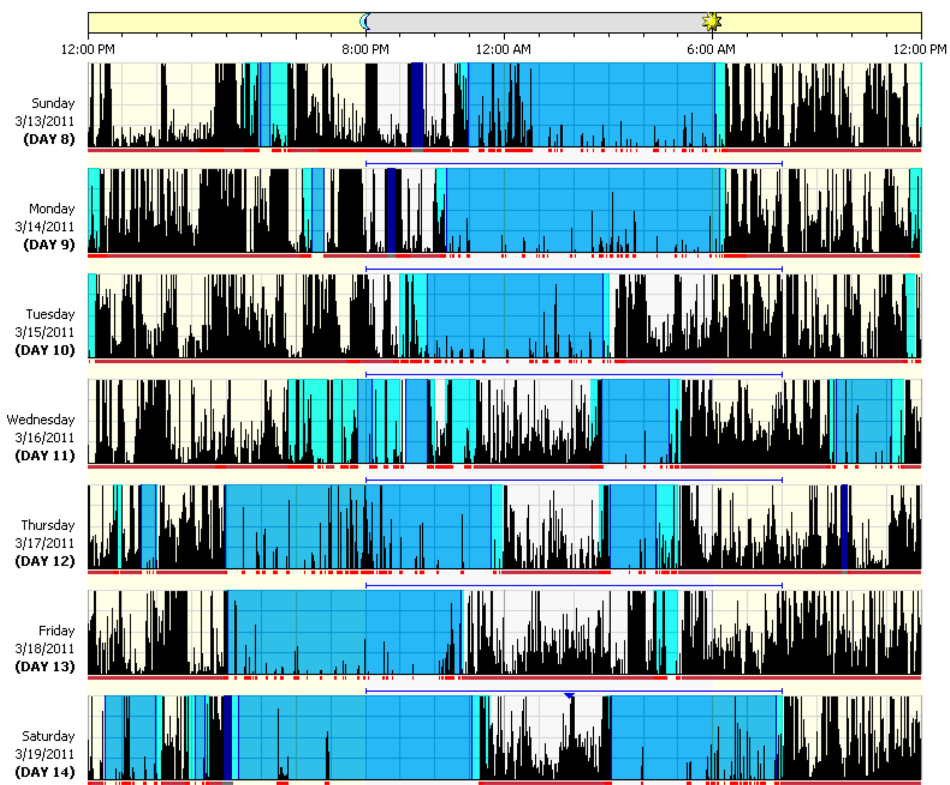


Figure 6. Example of actigraphy data imported from one actiwatch

Analysis was based on 15 participants, with an average of 8.6 days of actigraphy per person. Analysis was conducted on 129 days' worth of quality actigraphy data, including 52 days of inport data (40.3%), 57 of underway at low sea states (44.19%), and 20 of underway with high sea states (15.5%). The data was collected during eight days in port and nine days underway (two days included both in-port and underway data). On a daily basis, the number of participants ranged from five to thirteen with a mean of 9.21

and a standard deviation of 2.69 participants. Figure 7 depicts the number of participants per day. There was much lower participation during the inport period due to weekend and holiday ship routine.

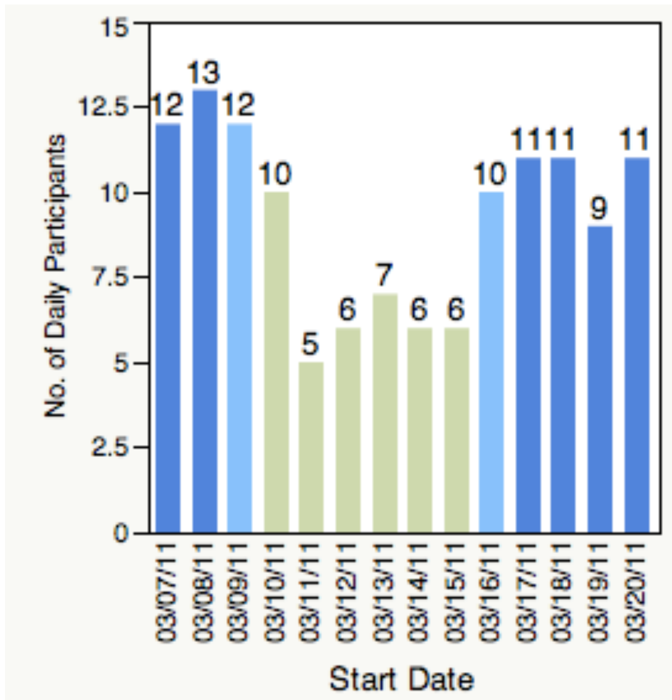


Figure 7. Actigraphy participation collected per day

3. FAST™ Export and Analysis Process

The next step in the data processing was to export it into the FAST™ software in order to determine predicted performance levels. The FAST™ user interface is a software package that takes work-sleep intervals and converts them to predicted effectiveness levels based on the SAFTE model. The FAST™ analysis can be conducted using manually inserted schedules, or by importing actigraphy data from the Actiware software. When importing the data, the sleep-wake information is displayed in one-minute epochs. For every minute, if activity is higher than 40 counts, then the epoch is considered "awake." If the activity count is less than 40, then the epoch is counted as "sleep" despite any activity intervals entered into Actiware. There are also smoothing options upon import that average the sleep-wake period every 5, 10, or 15 minutes. Prior to proceeding with the performance analysis, I wanted to determine the difference between the input options, and their effect on the accuracy of the output performance predictions.

Two participants with complete, uninterrupted actigraphy data throughout the entire testing period were evaluated in FAST^{TM} under multiple sleep input options: B572 and T313. Five different input options across the PVT performance test outputs were considered. The input options included straight actigraphy import without any alterations, actigraphy import with activity intervals (non-rest periods) smoothed to show as "awake," actigraphy import with the 5 minute smooth option, the manual sleep log inputs with environment set as "fair," and the pure manual sleep log inputs. Figures 8 and 9 show the FAST^{TM} imported data display for the two extreme choices.

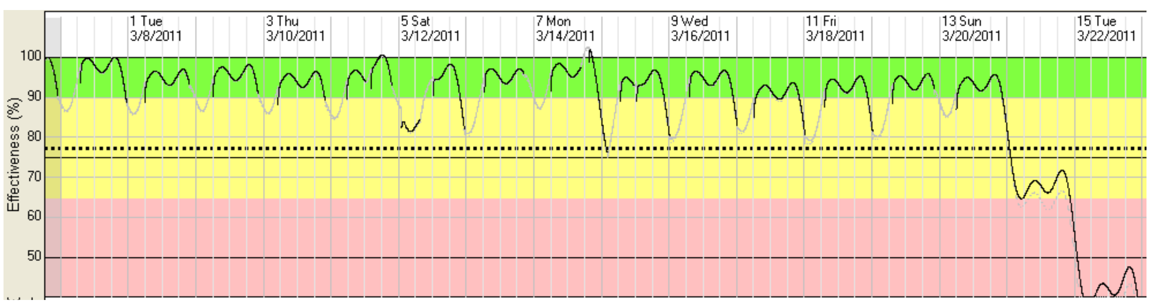


Figure 8. FAST^{TM} data display of pure actigraphy data import

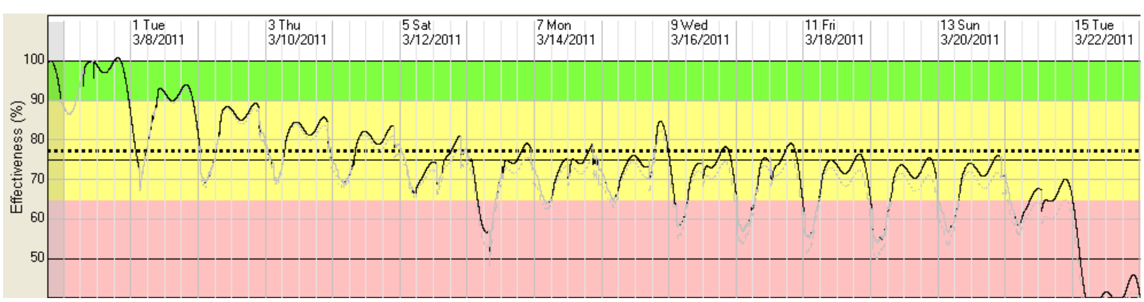


Figure 9. FAST^{TM} data display of manually imported self-reported activity log data

Table 1. Matched pairs t-test on FAST^{TM} predicted effectiveness outputs with the t statistic (t), p-values (p) and sample correlations (r)

N=45	Log Data w/Fair Env	Actigraphy w/5 min smooth	Actigraphy w/wake interval smooth	Pure Actigraphy
Pure Log Data	t=-22.66, p<0.0001** r=0.90	t=-14.22, p<0.0001** r=0.69	t=-20.00, p<0.0001** r=0.63	t=-19.96, p<0.0001** r=0.63
Pure Log Data w/Fair Env	-	t=-2.17, p=0.0356* r=0.62	t=-13.26, p<0.0001** r=0.71*	t=-12.20, p<0.0001** r=0.68
Actigraphy w/5 min smooth	-	-	t=-14.94, p<0.0001* r=0.85	t=-16.80, p<0.0001** r=0.90
Actigraphy w/wake interval smooth	-	-	-	t=3.82, p<0.0004** r=0.98

**refers to statistical significance at the 0.01 level (p<0.01)
* refers to statistical significance at the 0.05 level (p<0.05)

Paired t-tests and sample correlations in Table 1 show that although highly correlated, each FAST^{TM} models' mean predicted effectiveness score was significantly different than the rest.

The actigraphy data with wake intervals smoothed had the lowest predicted effectiveness, followed by the pure actigraphy data, then actigraphy data with 5 minute smoothing, followed closely by pure manual log data with fair environment setting and finally with the pure manually entered sleep log data. The actigraphy data allowed the SAFTE model to account for the disturbances during sleep that cause fragmentation and poor quality rest, whereas the manually entered log data did not. Figure 10 shows the predicted performance spectrum. The next step was to determine which model most accurately reflected the performance, and how the model could be adjusted to reflect the actual performance.

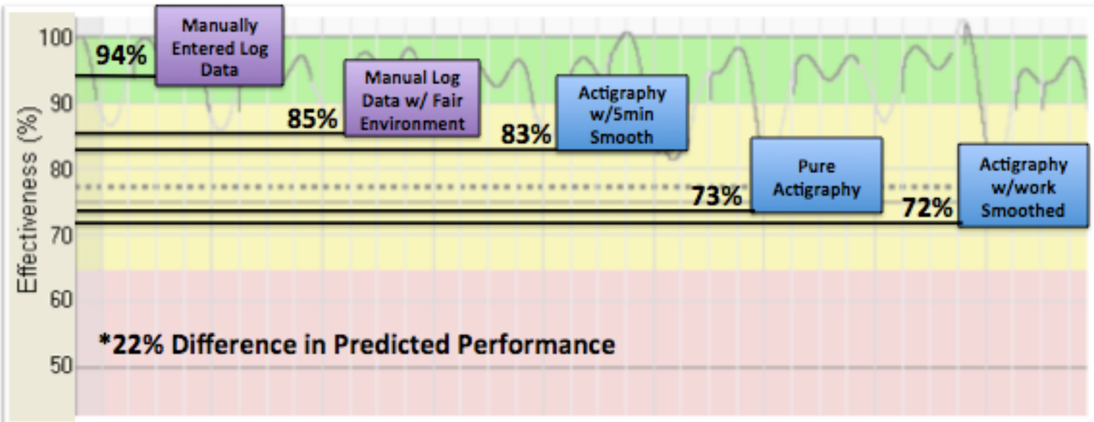


Figure 10. FAST™ output mean effectiveness (%) prediction spectrum

The predicted effectiveness levels were compared with the actual PVT performance levels. Since the SAFTE model does not account for the effects of maritime platforms, it was expected that the predicted effectiveness would be correlated with the actual performance only during the inport period. As seen in Table 2, only the FAST™ mean predictions extracted from the models based on actigraphy data differed from the mean PVT levels. The methodology of using actigraphy data with the wake intervals smoothed was chosen for the rest of the data exportation into the FAST™ program based on the results in Table 2.

Table 2. Paired t-test on performance models and actual PVT results for the inport period with the t statistic (t), p-values (p), and sample correlations (r)

N=20	Pure Log Data	Pure Log Data w/Fair Env	Actigraphy w/5 min smooth	Actigraphy w/wake interval smooth	Pure Actigraphy
PVT Mean RRT	t=0.148 p=0.8832 r=0.034	t=0.004 p=-.9600 r=0.012	t=2.175 p=0.0425* r=0.446	t=2.864 p=0.0099** r=0.549	t=2.614 p=0.0171* r=0.5143

**refers to statistical significance at the 0.01 level (p<0.01)
* refers to statistical significance at the 0.05 level (p<0.05)

The FAST™ interface relies on continuous sleep data for accurate predictions. For participants with incomplete actigraphy data (e.g., some excluded night sleeping period, or excluded entire 24-hour periods while in port), the sleep data had to be extracted from their activity log in order to maintain a continuous dataset. Figure 11 shows the number and quality of the actigraphy data sets per day. The participants with continuous data were designated as excellent (level “2”), and the participants whose actigraphy required supplemental data for some of the inport period were marked as adequate (level “1”). Participants who required supplemental data for the entire inport period or had both excluded actigraphy data and incomplete activity logs were marked a poor quality (level “0”). Participants designated a level “0” were not used in the analysis that followed.

Participant	3/7	3/8	3/9	3/10	3/11	3/12	3/13	3/14	3/15	3/16	3/17	3/18	3/19	3/20	3/21
A530	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A853	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0
B572	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
C169	2	2	1	2	2	0	2	2	0	2	2	2	0	2	2
C845	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D612	2	2	2	2	1	1	1	1	1	1	1	1	2	2	1
G700	2	2	2	2	1	1	1	0	0	0	0	0	0	0	0
i323	0	2	2	2	0	2	2	2	2	2	2	2	1	1	0
i499	2	2	0	0	0	0	2	2	2	2	2	2	2	2	2
K566	0	2	2	0	0	0	0	0	0	0	0	0	2	0	0
K597	0	2	2	1	1	1	1	1	1	1	2	2	1	2	1
K823	2	2	2	2	2	2	2	1	2	2	2	2	2	2	1
N364	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N446	0	1	1	1	1	1	1	1	1	2	2	2	2	2	2
R510	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
R572	2	2	0	0	2	2	2	2	2	2	2	2	2	2	2
T313	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
V516	2	2	2	2	1	0	1	0	0	2	2	2	2	1	0
X043	0	2	2	2	2	0	0	0	0	0	0	0	0	0	0
X866	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
Z772	2	2	0	2	0	1	0	0	0	2	2	2	1	2	1
Data Quality	50 %	70 %	50 %	50%	30 %	25%	35%	30%	30%	50%	55%	55%	45%	50%	30%

2	100% complete data quality dataset
1	Partial data (i.e., partial exclusion <24hr)
0	No activity for the day (i.e. 24 hr exclusion)

Figure 11. Quality of actigraphy data per day

4. Performance Tests

The data from performance tests were cleaned to remove any incomplete tests. Literature showed that the reciprocal transform (1/RT) of the mean reaction times has been proven to be sensitive to total and partial sleep loss (Basner & Dinges, 2011). This measure was used as the primary means of review.

The Switching Test showed a substantial learning curve as seen in Figure 12. In an attempt to prevent the learning effect from influencing the results, the first four tests per participant were excluded. Only test results of participants who had actigraphy data of quality "1" or "2" were used in the further analysis. There was no significant difference between the reaction time and reaction time for correct answers based on the paired t-test (p-value = 0.917). The overall reaction time was 5.4ms faster for correct answers. Figure 13 shows the average reaction times for both the Manikin and Math tests. There was some difference seen between the reaction time and reaction time for correct answers for the Manikin portion of the test based on the paired t-test (p-value = 0.094). The overall reaction time was 28.85ms faster for correct answers. There was a significant difference between the reaction time and reaction time for correct answers for the math cognitive test based on the paired t-test (p-value = 0.00005). The overall reaction time was 27ms faster for correct answers.

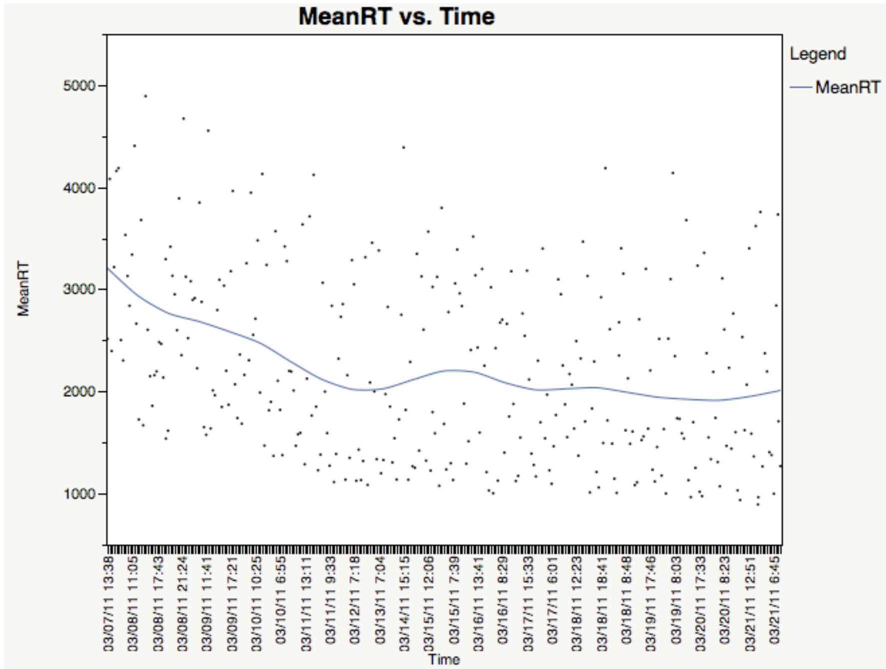
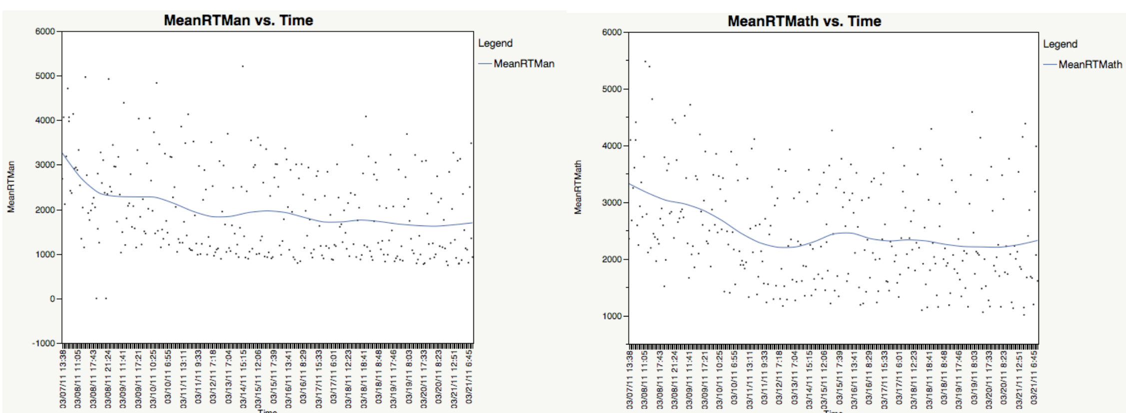


Figure 12. Switching test mean reaction time by date



(a) Manikin reaction time

(b) Math reaction time

Figure 13. Mean reaction time over time for a) Manikin test b) Math test

As expected from previous studies, the PVT did not show a learning effect (Figure 14). All of the PVT tests per participant whose actigraphy data quality was a “1” or “2” were used in the analysis.

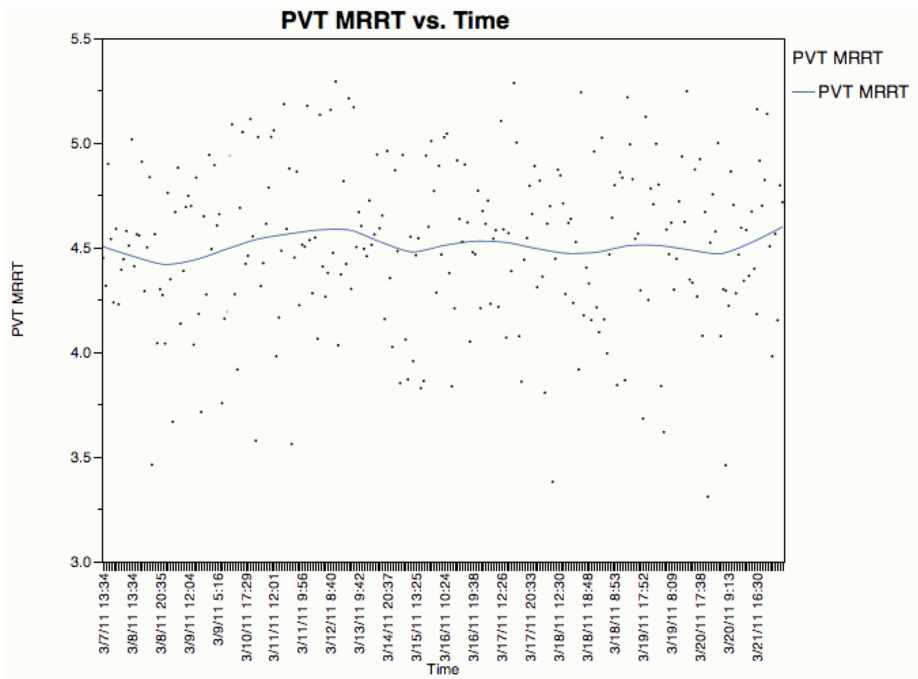


Figure 14. Psychomotor Vigilance Test (PVT) by date

THIS PAGE INTENTIONALLY LEFT BLANK

IV. RESULTS AND ANALYSIS

A. HYPOTHESES

The hypotheses are in two focus areas, sleep and performance. A clear relationship between sleep and sea state, or motion, has not been previously defined. It is predicted that increased motion during sleep periods will decrease the quantity and quality of sleep received by the sailors. It has also been shown that performance and coordination is directly impacted by physical motion (Wertheim, 1998); however, the focus of this thesis is on the relationship between motion, cognitive performance, and vigilance. Sleep directly impacts performance as found in the previous studies discussed in the literature review (Basner & Dinges, 2011; Belenky et al., 2003; Graw et al., 2001; Hursh et al., 2004; Johnson et al., 2004; Van Dongen et al., 2003). Consequently, it was expected that motion indirectly affects performance in a negative manner based on its direct effect on sleep quantity and quality.

Sleep Quantity: Sleep quantity decreases with increased ship motion as measured by sea state. The metrics used to determine sleep quantity are daily sleep duration in a 24-hour period and average sleep duration per sleep period.

Sleep Quality: Sleep quality decreases with increased ship motion as measured by sea state. The metrics used to determine sleep quality are the number of sleep bouts per sleep period, average activity count per sleep period, sleep efficiency per sleep period, and daily self-reported sleep problems.

Vigilance Performance: Vigilance performance decreases with the increase of ship motion as measured by sea state. The metrics used to determine vigilance are the mean reciprocal reaction time as measured through the three-minute PVT.

Cognitive Performance: Cognitive performance decreases with increased ship motion as measured by sea state. The metrics used to determine cognitive performance are the overall mean reaction time to respond to the Switching test and overall throughput for the ANAM Switching test. The throughput is measured as the number of correct responses per minute.

With the numerous sources of data, the multivariate nature of the responses, and the issues of consistent subject compliance, the analysis presented investigated most of the hypotheses using many simple procedures rather than using a few inappropriate and unnecessarily complex models. The danger of finding patterns when none exists comes with analysis using many hypothesis tests. To mitigate the increased family-wise probability of type one error with multiple tests, a Bonferroni adjustment was applied and only tests with p-values less than .001 were highlighted as statistically significant.

Based on the results from testing these basic sleep and performance hypotheses, this study determined the adequacy of using the SAFTE model used in the FAST™ interface for maritime platforms, based on its ability to account for the direct effects of motion on sleep and thereby having an indirect effect on performance outcome. It was predicted that the FAST™ predicted effectiveness performance levels would be less accurate as sea state increases.

B. SLEEP DATA

Sleep data were collected in both an objective and subjective manner. Seven sleep-related metrics were derived from the actigraphy data (rest time, average activity per minute, sleep efficiency, percent awake time, sleep time, percentage of sleep time, and average number of sleep bouts per minute). Daily sleep ratio, or number of sleep episodes per day, was also derived. Sleep quality was also measured through the subjective evaluations provided by the NATO PAQ "Sleeping problems" group of questions as described in chapter three and the daily activity logs.

1. Participant Descriptive Statistics

Sleep data were collected from 21 crewmembers. From that group, data from only 15 participants were usable (14 males, 1 female). The ranks of the participating crewmembers are depicted in Figure 15. The mean age of the participants was 35.8 years with a standard deviation of 5.92 years. With a single exception, all participants scored as a "good" or higher on their most recent Naval physical readiness test.

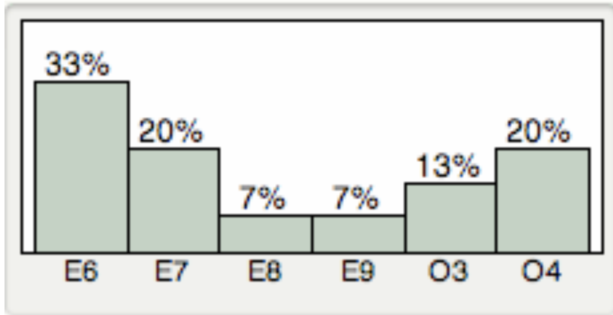


Figure 15. Rank of participants

Watch rotation among the participants varied. Nine participants (60%) were on set watch shifts or did not stand watch while underway while six (40%) were on various rotating watch schedules.

When discussing sleep effects on performance, it is important to take into account factors that may affect sleep quality and enhance performance such as caffeine intake, tobacco use, and use of sea sickness medication. In the background survey, the participants were asked about each of these categories. Seven of the fifteen participants used tobacco products (47%). Seven participants (47%) reported using seasickness medication while at sea, an important factor to consider since many types of seasickness medication cause sleepiness. Finally, 60% of the participants reported requiring one or more caffeinated beverages a day (coffee, soda, or energy drink). Figure 16 shows the distribution of caffeine intake per day.

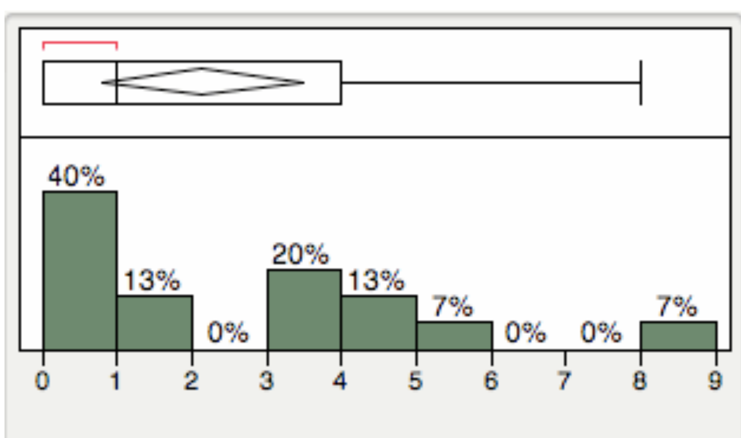


Figure 16. Caffeinated beverage intake per day

2. Actigraphic Data

Actigraphic data were collected for 21 participants. The quality of the data collected varied, due to actiwatch device malfunctions as well as compliance issues. For example, some participants took the device off during the nighttime sleep period, while others took it off when on liberty during the inport period. Not surprisingly, the quality of data varied based on these compliance challenges. Figure 11 in the previous section showed the distribution of the data quality per participant throughout the collection period. Based on the data quality, the sleep analysis included only 15 participants who are indicated in green.

Next, sleep analysis was performed. Table 3 shows the sleep summary statistics from the actigraphy data, comparing the inport versus underway metrics using a two-sample t-test. In the analysis, the variability between participants was isolated by blocking on individual, therefore more accurately reflecting the differences in the sleep data due to the change in motion condition.

Table 3. Actigraphy determined sleep metrics by ship status (blocked on participant)

Metric	Inport Mean (StdDev)	Underway Mean (StdDev)	t-stat (df) (Prob> |t|)
Daily Wake Time (%) (24hr)	70.28(10.67)	68.14(12.02)	t(117)= -1.736 p=0.0852
Daily Sleep Time (hrs) (24hr)	6.81(2.31)	7.36(2.63)	t(117)=2.075 p=0.0402
Average Daily Sleep Bouts (mins/bout) (24hr)	9.15(5.53)	10.69(6.26)	t(114)=2.584 p=0.0110
Sleep Time (hrs) per Sleep Episode	3.68(2.70)	3.93(3.03)	t(155)= 0.881 p=0.3795
Avg Activity during Sleep (counts/min) per Sleep Episode	21.13(26.06)	26.24(28.98)	t(152)= 1.889 p=0.0607
Sleep Efficiency (%) per Sleep Episode	60.75(24.4)	61.68(27.28)	t(152)= 0.390 p=0.697
Average Sleep Bouts (mins/bout) per Sleep Episode	13.52(35.40)	18.42(35.32)	t(159)= 0.924 p=0.3567

Daily inport N=52, Daily underway N=77 ; Sleep inport N=64, Sleep underway N=99

Although participants had higher mean daily sleep time during the underway period, the actigraphic results indicate that the sleep was of poor quality. The higher average activity counts per minute during sleep periods for the underway periods is an indicator of poor sleep quality. These findings are consistent with the self-reported data shown in the next section.

Thus far, analysis of sleep disturbances was based on the comparison between inport and underway conditions. Given that motion is one of the major differences between the inport and underway stressors, this comparison is used as a baseline in order to assess the effect of motion of sleep. Yet, sleep disturbances are caused by numerous factors, some of which exist both inport and underway. Therefore, it is logical to expect that the extent of such problems while underway is partially attributed to factors other than the existence of motion. The next step was to evaluate the association between the severity of sleep disturbances and sea state. The analysis summarized in Table 4 was based on the amount of motion or sea state (SS) and divided into two groups (Low/High).

Table 4. Actigraphy determined sleep metrics by sea state

Metric	Low SS Mean (StdDev)	High SS Mean (StdDev)	t-stat(df) (Prob>|t|)
Daily Wake Time (%) (24hr)	68.31(12.68)	66.48(9.06)	t(63)= -1.145, p=0.2565
Daily Sleep Time (hrs) (24hr)	7.21(2.91)	7.97(2.00)	t(63)=2.356, p=0.0216
Daily Sleep Time (%) (24hr)	31.70(12.68)	33.52(9.06)	t(63)= 1.145, p=0.2567
Average Daily Sleep Bouts (mins/bout)	10.50(6.11)	11.53(4.92)	t(62)=0.978, p=0.3319
Sleep Time (hrs) per Sleep Episode	4.01(3.13)	3.70(2.38)	t(88)= -0.791, p=0.4310
Avg Activity during Sleep (counts/min) per Sleep Episode	24.77(30.73)	30.11(22.80)	t(84)= 1.449, p=0.1509
Sleep Efficiency (%) per Sleep Episode	60.72(24.83)	60.43(19.00)	t(87)= -0.093, p=0.9261
Average Sleep Bouts (mins/bout) per Sleep Episode	17.42(44.80)	20.86(42.51)	t(89)= 0.379, p=0.7059

Daily LSS N=57, Daily HSS N=20 ; Sleep LSS N=72, Sleep HSS N=27

Based on the results, the duration of the sleep episodes while in port and underway at the sea states were further evaluated. The analysis (Figure 17) suggests that there were many more sleep episodes per day while underway as compared to the inport sleep episodes (underway: Mean=14.00 episodes/day, StdDev=4.56; inport: Mean=9.88 episodes/day, StdDev=3.48; $t(12)=1.925$, $p=0.0782$).

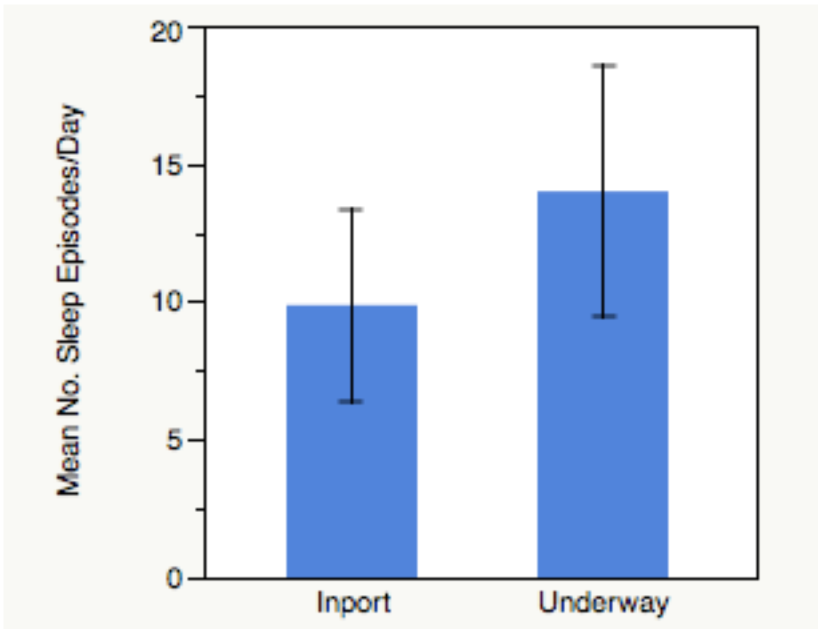


Figure 17. Mean number of sleep episodes per day by ship status with standard deviation bars

3. Questionnaire Responses

The next step in the analysis was to address the information provided in the subjective test questionnaires. Table 5 demonstrates the findings by comparing the inport versus the underway conditions using the Wilcoxon Rank Sum Test. As expected, more sleep problems were reported during the underway period, and of those problems, a significant proportion was attributed to the ship's motion.

Table 5. Sleep problems responses by motion condition using Wilcoxen Rank Sum test

Questionnaire Statement Inport N=51, Underway N=78	Inport Mean (StdDev)	Underway Mean (StdDev)	Wilcoxen Rank Sum stat P-Value
Quality of sleep was poor	0.412 (0.497)	0.544 (0.501)	W =2.16, p=0.1415
Amount of time sleeping was short	0.538 (0.503)	0.61 (0.49)	W =0.6547, p=0.4184
Sleep problems were caused by ship motion	0.118 (0.325)	0.462 (0.502)	W=16.48, p<0.0001*
Sleep problems were caused by seasickness	0.039(0.2)	0.115(0.32)	W =2.2759, p=0.1314
Sleep problems were caused by other factors	0.12 (0.32)	0.14 (0.35)	W =0.1956, p=0.658
Overall sleep problems - total score (Inport N=256; Underway N=395)	0.246(0.43)	0.37(0.48)	W =11.28, p<0.0008*

* refers to statistical significance at the 0.001 level (p<0.001).

Next, fatigue related to sleep time and sleep quality was assessed. Sleep metrics were extracted from actigraphy data through the FAST™ interface (sleep received in the last 24 hours, reservoir level, time awake, and chronic sleep debt). The subjective fatigue evaluations were extracted from the NATO PAQ questionnaires. Participants were asked to rate their severity of mental fatigue, physical fatigue, and sleepiness on a 4-point Likert scale (0=not at all, 3=extreme). In addition, subjective evaluation of sleepiness was measured using the Stanford Sleepiness Scale (SSS). All analyses regarding fatigue versus sleep attributes were conducted in two ways. First, a comparison between the inport versus the underway conditions was conducted. Next, a comparison between the sea states of the sleep information taken from the actigraphy data through the FAST™ output during the times that the participants took the NATO PAQ and SSS questionnaires was conducted.

In the analysis of inport versus underway conditions, no difference was found in levels of fatigue reported. Next, the relationship between the sleep metrics and reported fatigue and sleepiness were examined. The results are shown in Table 6. As expected, those individuals with higher chronic sleep debt and lower sleep reservoir levels reported increased sleepiness. Additionally, those participants with fewer hours of sleep in the last 24 hours reported increased sleepiness.

Table 6. Relationship between the mean sleep metrics and reported fatigue or sleepiness (blocked on participant)

Questionnaire Statement	No Fatigue Mean (StdDev)	Fatigue Reported Mean (StdDev)	t-stat(df) (Prob>|t|)
Chronic Sleep Debt			
Sleepiness (NATO PAQ)	5.11(3.4)	6.6(4.3)	t(115)=2.17, p=0.0321
Sleep Last 24hrs			
Sleepiness (SSS)	7.17(1.65)	6.64(1.57)	t(115)= -1.76, p=0.081
Hours Awake			
Mental Fatigue	4.83(5.54)	3.23(3.33)	t(115)= -1.81, p=0.0729
Physical Fatigue	4.9(5.4)	3.0(3.25)	t(115)= -2.15, p=-0.0339
Sleepiness (NATO PAQ)	5.75(5.8)	3.22(3.7)	t(115)= -2.81, p=0.0058
Reservoir Level			
Sleepiness (NATO PAQ)	84%(10%)	79%(10%)	t(115)= -2.36, p=0.0198

Finally, the relationship between the sleep metrics, reported fatigue, and the sea states was examined. The results in Table 7 indicates a pattern showing differences in the amount of sleep received over the last 24 hours in participants who report fatigue and sleepiness. The underway periods showed higher numbers of hours of sleep than the inport period, indicating reduced sleep quality while underway.

Table 7. Relationship between the average amount of sleep received in previous 24 hours and fatigue, by sea state

Questionnaire Statement or Metric	Inport Mean (StdDev)	UW – LSS Mean (StdDev)	UW – HSS Mean (StdDev)	F-Ratio(df) P-Value	Tukey Test HSD
Sleep Last 24hrs					
Mental Fatigue (No: 0)	7.21(1.4)	6.63(1.33)	6.84(1.8)	F(2,61)=0.9443, p=0.3946	-
Mental Fatigue (Yes: 1–3)	6.11(1.56)	7.00(1.76)	7.78(1.79)	F(2,48)=3.76, p=0.0304	HSS vs. Inport Diff=1.66, p=0.0260
Physical Fatigue (No: 0)	7.06(1.47)	6.54(1.4)	6.52(1.27)	F(2,65)=1.177, p=0.3145	
Physical Fatigue (Yes: 1–3)	6.21(1.59)	7.1(1.66)	8.54(2.01)	F(2,44)=5.912, p=0.0053	HSS vs. Inport Diff=2.32, p=0.0036 HSS vs. LSS Diff=1.43, p=0.0932
Sleepiness (SSS) (No: 1)	7.35(1.56)	7.06(1.92)	7.06(1.5)	F(2,48)=0.174, p=-.8408	
Sleepiness (SSS) (Yes: 2–7)	6.26(1.42)	6.61(1.21)	7.44(2.19)	F(2,61)=2.56, p=0.0853	HSS vs. Inport Diff=1.18, p=0.0691

To investigate this further, the difference within each sea state was evaluated. As seen in Table 8, the reported fatigue only had differences in hours of sleep and reservoir levels inport, while underway sleep estimates were similar between participants who reported fatigue and those who did not. This indicates that additional factors associated with the motion of being underway may be attributed to the feelings of fatigue, and not solely the amount of sleep received.

Table 8. Relationship between mean sleep metrics and fatigue, by sea state (blocked on participant)

Questionnaire Statement or Metric	No Fatigue	Fatigue Reported	t-stat(df) (Prob>|t|)
Mental Fatigue			
Sleep Last 24hrs	**Mean (StdDev)**	**Mean (StdDev)**	
Inport	7.2(1.4)	6.11(1.5)	t(45)= -2.48, p=0.0171
Underway - LSS	6.63(1.33)	7.00(1.76)	t(42)=0.77, p=0.4455
Underway - HSS	6.84(1.8)	7.78(1.79)	t(28)=1.37, p=0.1830
Reservoir Level			
Inport	85%(2%)	78%(2%)	t(45)= -2.2, p=0.0331
Underway - LSS	80%(11%)	78%(11%)	t(42)=0.63, p=0.6303
Underway - HSS	80%(12%)	80%(10%)	t(28)= -0.10, p=0.9234
Physical Fatigue			
Sleep Last 24hrs	**Mean (StdDev)**	**Mean (StdDev)**	
Inport	7.06(1.47)	6.21(1.59)	t(45)= -1.83, p=0.0743
Underway - LSS	6.54(1.4)	7.11(1.66)	t(42)=1.20, p=0.2387
Underway - HSS	6.52(1.27)	8.54(2.01)	t(28)=3.26, p=0.0031
Reservoir Level			
Inport	84%(10%)	78%(9%)	t(45)= -2.15, p=0.0369
Underway - LSS	78%(13%)	80%(8%)	t(42)=0.48, p=0.4813
Underway - HSS	80%(12%)	81%(10%)	t(28)= 0.25, p=0.8039
Sleepiness (SSS)			
Sleep Last 24hrs	**Mean (StdDev)**	**Mean (StdDev)**	
Inport	7.35(1.56)	6.26(1.42)	t(45)= -2.43, p=0.0196
Underway - LSS	7.06(1.92)	6.61(1.21)	t(42)= -0.94, p=0.3534
Underway - HSS	7.06(1.5)	7.44(2.19)	t(28)= 0.54, p=0.5926

Table 9. Correlations between fatigue, sleepiness, and sleep received in the previous 24 hours, hours awake, and sleep reservoir, by sea state

Performance Metric	Spearman Correlation
Inport	
Mental Fatigue x Sleep	r_s=-0.2906, p=0.0528
Mental Fatigue x Reservoir	r_s=-0.3488, p=0.0188
Physical Fatigue x Reservoir	r_s=-0.3264, p=0.0286
Sleepiness(SSS) x Sleep	r_s=-0.3189, p=0.0327
Sleepiness(SSS) x Hours Awake	r_s=-0.2564, p=0.0891
Underway – Low Sea State	
Physical Fatigue x Hours Awake	r_s=-0.3101, p=0.0456
Underway – High Sea State	
Physical Fatigue x Sleep	r_s=0.4542, p=0.0152
Sleepiness (PAQ) x Reservoir	r_s=-0.3650, p=0.0561

The findings suggest that fatigue reported in the inport period could be attributed to actual reduced sleep reservoir levels and sleep conditions. The fatigue reported during the underway period could be due to motion related symptoms (sopite syndrome) and deteriorated sleep quality rather than reduced sleep quantity. Increased fatigue levels were seen underway compared to inport conditions. Increased fatigue levels reported while inport were significantly associated with 15% fewer hours of sleep in the previous 24 hours. Although crewmembers associated fatigue and sleepiness with deteriorated sleep quantity and quality, these findings could not be explained by the objective sleep metrics measuring sleep quantity during the underway period.

4. Sleep Regression Analysis

Through a process of trial and error, various models were explored to show the motion effects on sleep quantity and quality at sea. Because of the significant within subjects variability among the participants, a mixed effects model was used whereby individual differences were treated as a random effect and the other variables were treated as fixed effects. The independence, equal variance, and normality assumptions and conditions were met when using the mixed effects model. All sleep models in this thesis used the mixed effects model with the participant as the random effect, producing similar diagnostic plots, indicating that the modeling assumptions are met for each model.

Many models were fit when assessing the sleep data. These models included independent variables such as sea state, shift type, rank, group type, and their interactions. These variables were included due to their potential direct effect on schedule, work habit, motivation, and therefore, sleep cycles. Although many combinations proved significant, the sleep models for Daily Sleep Quantity, Sleep Quantity per Sleep Period, and Sleep Quality with the best fits based on the data collected are summarized in Tables 10, 11, and 12.

Table 10. Daily sleep quantity random effects model summary with F-test statistic to test for effects

R^2= 0.50 R^2_{adj}=0.46	**Coefficient Estimates (SE)**	**F-Ratio (df)**	**P-Value**
Seas	Inport: -0.543 (.36) LSS: 0.305 (.34) HSS: 0.238 (.49)	F(2,111)=3.33	0.0391*
Rank	Jr. Enlisted: 0.72 (.50) Sr. Enlisted: 0.36 (.42) Officer: -1.08 (.47)	F(2, 10)=5.27	0.0272*
Work Group	Maintainer: 0.150 (.60) Operator: 0.204 (.29) Sr. Leadership: -0.353 (.67)	F(2,12)=0.29	0.7548
Work Group x Seas	Maintainer x Inport: -1.103 (.69) Maintainer x HSS: 1.058 (.92) Operator x LSS: -0.726 (.32) Operator x HSS: 0.526 (.43) Sr Leadership x Inport: 0.9 (.78) Sr Leadership x LSS: 0.68 (.71) Sr Leadership x HSS: -1.58 (1.08)	F(4, 111)=4.73	0.0015*

Table 11. Sleep quantity per sleep period random effects model summary with F-test statistic to test for effects

R^2=0.17 , R^2_{adj}=0.16	Coefficient Estimates (SE)	F-Ratio (df)	P-Value
Work Group	Maintainer: 1.20 (0.71) Operator: -1.03 (0.33) Sr. Leadership: -0.171 (0.75)	F(2, 17)=4.210	0.0323*

Table 12. Sleep quality random effects model summary with F-test statistic to test for effects

a) Number of sleep bouts during sleep session

R^2=0.29 , R^2_{adj}=0.255	Coefficient Estimates (SE)	F-Ratio (df)	P-Value
Seas	Inport: 1.15 (2.52) LSS: 4.99 (2.11) HSS: -6.14 (3.56)	F(2,148)=4.020	0.020*
Work Group	Maintainer: 10.48 (3.48) Operator: -5.20 (1.5) Sr. Leadership: -5.28 (3.59)	F(2,16)=8.743	0.003*
Work Group x Seas	Maintainer x Inport: -3.55 (4.65) Maintainer x LSS: 7.05 (4.31) Maintainer x HSS: -3.50 (7.3) Operator x Inport: -1.61 (2.0) Operator x LSS: -4.72 (1.91) Operator x HSS: 6.34 (2.84) Sr Ldrship x Inport: 5.16 (5.65) Sr Ldrship x LSS: -2.32 (4.23) Sr Ldrship x HSS: -2.84 (7.27)	F(4,146)=2.596	0.039*

b) Average activity count per minute during sleep session

R^2=0.289 , R^2_{adj}=0.28	Coefficient Estimates (SE)	F-Ratio (df)	P-Value
Seas	Inport: -4.368 (3.23) LSS: -0.779 (3.07) HSS: 5.147 (4.01)	F(2,150)=3.048	0.050*

These models show that the sleep quantity actually increased with the increase in ship motion or sea state. However, sleep quality decreased with the increase in motion as seen with the increase in activity per sleep period. The decrease in number of sleep bouts per sleep period with the increase in motion may indicate changes in sleep architecture.

There also appeared to be no change in sleep efficiency across sea state (model not significant and not included in Table 11 or 12).

C. PERFORMANCE DATA

Performance data was collected through the use of two performance tests as discussed in Chapter Three. The goal of this section is to show that both vigilance and cognitive performance decrease with the increase of motion as measured by sea state. This section will then compare the differences between performance as predicted by the SAFTE model through the FAST™ interface and the actual performance as measured by the performance tests. Based on these results, potential improvements to the SAFTE model will be discussed in the next section.

1. Actual Performance by Sea State

a. Vigilance Performance

Vigilance performance was measured through the mean reciprocal reaction time (RRT), median RRT, fastest RRT, and slowest RRT outputs using the PVT.

Table 13. Vigilance by ship status (two-sample t-test with block on participant)

Metric (PVT)	Inport N=100 Mean (StdDev)	Underway N=166 Mean (StdDev)	t-stat(df) (Prob> |t|)
Mean RRT	4.54(0.80)	4.47(1.00)	t(266)= -2.12, p=0.0350
Median RRT	4.64(0.88)	4.58(1.10)	t(266)= -1.80, p=0.0723
Fastest RRT	5.46(0.67)	5.49(0.81)	t(266)= 0.757, p=0.4496
Slowest RRT	3.17(0.93)	3.07(1.08)	t(266)= -1.517, p=0.1304

Table 14. Vigilance performance by ship motion (two-sample t-test, blocked on participant)

Metric (PVT)	Inport N=102 Mean (StdDev)	Underway N=64 Mean (StdDev)	t-stat(df) (Prob> |t|)
Mean RRT	4.45(0.80)	4.51(0.66)	t(166)= 1.23, p=0.2205
Median RRT	4.57(0.90)	4.59(0.74)	t(166)= 0.58, p=0.5616
Fastest RRT	5.49(0.64)	5.47(0.54)	t(166)= -0.48, p=0.6337
Slowest RRT	3.01(0.81)	3.17(0.74)	t(166)= 1.83, p=0.0687

There was a possible pattern indicating an increase in vigilance performance across both mean RRT and median RRT measures between inport and underway as seen in Table 13. The slowest RRT measure showed the only indication of a relationship with the change in sea state as shown in Table 14.

b. *Actual Cognitive Performance*

Cognitive performance was measured by the overall mean reaction time to respond to the task, overall throughput, throughput for Manikin test, and throughput for Math computation. The throughput is measured as the number of correct responses per minute.

Table 15. Cognitive performance by ship status (two-sample t-test with blocked on participant)

Metric	InPort N=98 Mean (StdDev)	Underway N=98 Mean (StdDev)	t-stat(df) (Prob> |t|)
MRRT - Overall	4.87(5.84)	5.69(5.84)	t(196)= 7.61, p<0.0001*
MRRT - Correct	5.70(5.88)	4.90(5.86)	t(196)= 7.35, p<0.0001*
Throughput - Overall	26.92(38.57)	31.93(38.47)	t(196)= 7.35, p<0.0001*
Throughput - Manikin	30.7(55.38)	39.0(55.32)	t(196)= 9.74, p<0.0001*
Throughput - Math	24.82(28.15)	27.84(28.08)	t(196)= 5.29, p<0.0001*

* refers to statistical significance at the 0.001 level (p<0.001)

Table 16. Cognitive performance by ship motion (two-sample t-test with blocked on participant)

Metric	Low SS N=45 Mean (StdDev)	High SS N=53 Mean (StdDev)	t-stat(df) (Prob> |t|)
MRRT - Overall	5.46(4.25)	5.84(4.61)	t(98)= 3.17, p=0.0021
MRRT - Correct	5.47(4.26)	5.84(4.62)	t(98)= 2.93, p=0.0043
Throughput - Overall	31.01(26.83)	32.7(29.05)	t(98)= 2.76, p=0.0072
Throughput - Manikin	36.98(38.17)	40.45(41.35)	t(98)= 3.92, p=0.0002*
Throughput - Math	27.35(20.12)	27.87(21.84)	t(98)= 0.75, p=0.4542

* refers to statistical significance at the 0.001 level (p<0.001)

There was a definite significant increase in cognitive performance as measured by manikin test throughput, but indications of improvements across all measures between inport and underway and between sea states with the exception of math throughput rate as seen in Tables 15 and 16. This may indicate that there was a continuous learning effect throughout the trials. Figure 18 further explores the possible learning effect across test number for both the Switching and PVT. There are indications that learning continued in the Switching Test through trial number 13, which was beyond the last test administered for many participants. The PVT continued to show no learning effect.

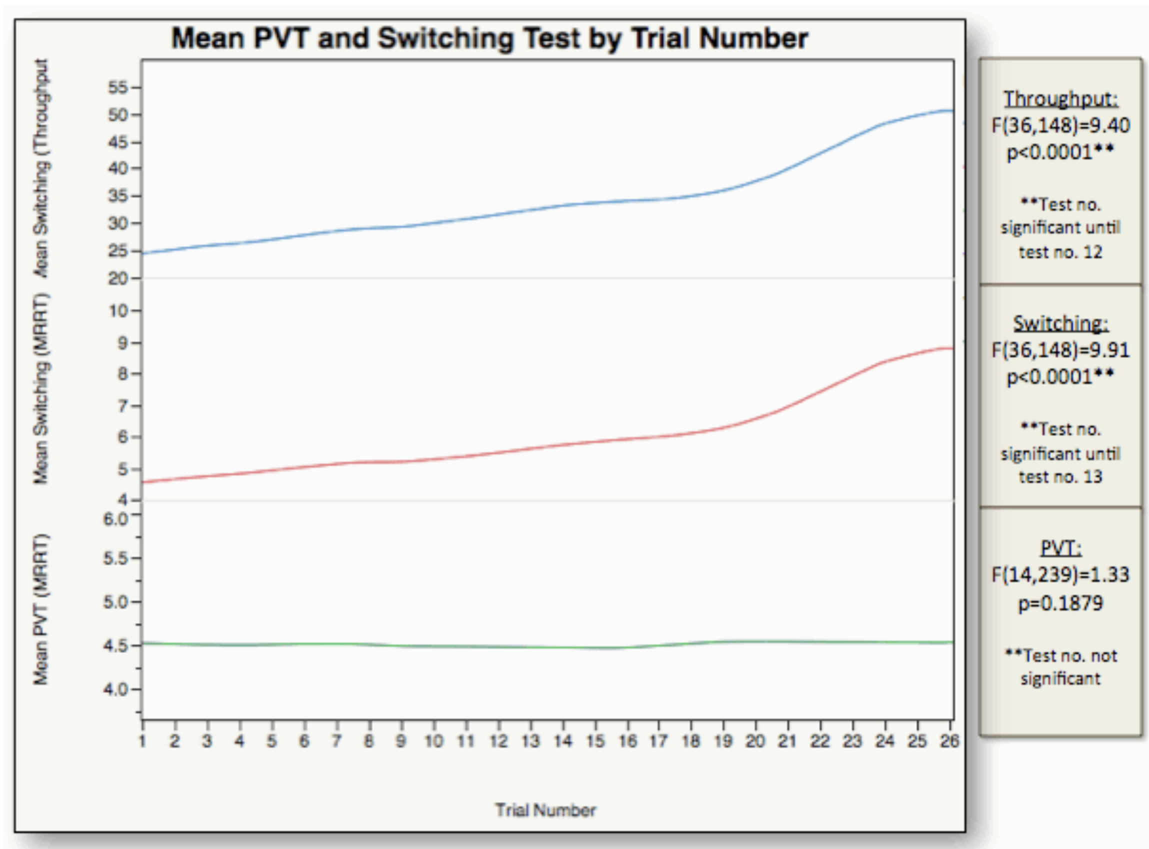


Figure 18. Mean performance across trial number for Switching test and PVT (blocked on participant)

Table 17 is a summary of the cognitive performance measures categorized by the participants' reservoir level at the time of the test. Even with the apparent learning effect, the analysis showed significantly higher performance at higher reservoir levels during the inport period, but not during the underway period. This finding may be due to the increased fragmentation of sleep observed during the underway. There is also much higher variability in sleep quantity and quality during the underway period.

Table 17. Relationship between cognitive performance and sleep reservoir level, by ship status

PerformanceMetric	Low (<65%) Mean (StdDev)	Medium (65–90%) Mean (StdDev)	High (>90%) Mean (StdDev)	F-Ratio(df) P-Value	Tukey Test HSD
Inport					
MRRT Overall	4.65(1.85)	5.41(2.22)	7.15(1.48)	F(2,94)=6.52, p=0.0022	Hi vs. Low Diff=2.5, p=0.0099 Hi vs. Med Diff=1.74, p=0.0045
Throughput Overall	26.4(12.6)	30.0(14.3)	41.3(8.56)	F(2,94)=6.17, p=0.0030	Hi vs. Low Diff=14.86, p=0.0186 Hi vs. Med Diff=11.22, p=0.0043
Throughput Manikin	29.7(19.0)	34.4(19.8)	52.2(11.7)	F(2,94)=7.79, p=0.0007*	Hi vs. Low Diff=22.49, p=0.0094 Hi vs. Med Diff=17.81 p=0.0010*
Throughput Math	24.7(8.53)	27.5(11.1)	34.2(7.1)	F(2,94)=3.89, p=0.238	Hi vs. Low Diff=9.48, p=0.0614 Hi vs. Med Diff=6.68, p=0.0352
Underway					
MRRT Overall	5.95(0.94)	6.01(2.66)	5.74(2.99)	F(2,95)=0.071, p=0.9311	-
Throughput Overall	34.97(5.55)	33.47(15.5)	31.5(20.7)	F(2,95)=0.213, p=0.8080	
Throughput Manikin	44.61(7.7)	40.74(22.1)	36.09(27.35)	F(2,95)=0.657, p=0.5206	
Throughput Math	28.9(5.21)	28.98(11.9)	28.5(16.1)	F(2,95)=0.011, p=0.9889	

* refers to statistical significance at the 0.001 level (p<0.001)

There was positive significant correlation between participant sleep reservoir level and performance as scored by MRRT and all throughput types during the inport period. However, this correlation was not seen while underway, indicating that other factors underway may override the effect of reservoir level on cognitive performance. Table 18 shows the correlation.

Table 18. Correlation between cognitive performance and reservoir level, by sea state

Performance Metric	Spearman Correlation
Inport	
MRRT Overall x Reservoir	r_s=0.475, p<0.0001*
Throughput Overall x Reservoir	r_s=0.468, p<0.0001*
Throughput Manikin x Reservoir	r_s=0.455, p<0.0001*
Throughput Math x Reservoir	r_s=0.434, p<0.0001*
Underway – Low Sea State	
MRRT Overall x Reservoir	r_s=0.049, p=0.7509
Throughput Overall x Reservoir	r_s=0.034, p=0.8264
Throughput Manikin x Reservoir	r_s=0.039, p=0.7974
Throughput Math x Reservoir	r_s=-0.013, p=0.9335
Underway – High Sea State	
MRRT Overall x Reservoir	r_s=-0.028, p=0.8411
Throughput Overall x Reservoir	r_s=-0.003, p=0.9831
Throughput Manikin x Reservoir	r_s=-0.084, p=0.5508
Throughput Math x Reservoir	r_s=0.018, p=0.8963

* refers to statistical significance at the 0.001 level (p<0.001)

Next, the differences in performance within each reservoir level across sea state were explored. As shown in Table 19, there is little difference in performance across the sea state levels at high and medium reservoir level (with the exception of the Throughput for the Manikin test). However, at low sleep reservoir levels, there are indications of possible differences. The performance actually increased underway. Since the inport period was still within the early phase of test taking, this may indicate that there is a longer learning curve effect for the cognitive testing in participants at low reservoir sleep conditions.

Table 19. Relationship between cognitive performance and sea state, by sleep reservoir level

Sleep Metric	Inport	Underway LSS	Underway HSS	F-Ratio(df) P-Value	Tukey Test HSD
High Reservoir Level (>90%)	**Mean (StdDev)**	**Mean (StdDev)**	**Mean (StdDev)**		
MRRT Overall	7.65(1.4)	6.11(2.75)	5.98(3.7)	F(2,32)=1.96, p=0.1574	
Throughput Overall	41.27(8.6)	32.99(18.3)	29.98(24.0)	F(2,32)=1.80, p=0.1810	
Throughput Manikin	52.25(11.74)	39.2(25.05)	32.99(30.9)	F(2,32)=2.86, p=0.0719	IP vs. UW HSS Diff=19.26, p=0.0828
Throughput Math	24.18(7.1)	28.95(13.9)	28.01(19.1)	F(2,32)=0.94, p=0.3974	
Medium Reservoir Level (65–90%)					
MRRT Overall	5.87(2.31)	6.07(2.64)	6.86(3.05)	F(2,131)=1.73, p=0.1801	
Throughput Overall	30.04(14.34)	30.6(15.06)	35.8(2.62)	F(2,131)=1.85, p=0.1600	
Throughput Manikin	34.43(19.8)	36.4(21.29)	44.2(22.38)	F(2,131)=2.65, p=0.0743	UW-HSS vs. IP Diff=9.78, p=0.0620
Throughput Math	27.5(11.05)	27.2(11.76)	30.4(26.27)	F(2,131)=0.88, p=0.4172	
Low Reservoir Level (<65%)					
MRRT Overall	5.01(1.82)	6.54(0.97)	6.5(1.3)	F(2,23)=3.3, p=0.0545	UW-LSS vs. IP Diff=1.53, p=0.0.0907 UW HSS vs. IP Diff=1.49, p=0.0881
Throughput Overall	26.4(12.61)	36.3(4.83)	33.8(6.14)	F(2,23)=3.03, p=0.0679	UW-LSS vs. IP Diff=9.89, p=0.0.0697
Throughput Manikin	29.8(19.03)	44.63(7.8)	44.6(8.08)	F(2, 23)=3.87, p=0.0354	UW-LSS vs. IP Diff=14.87, p=0.0664 UW HSS vs. IP Diff=14.83, p=0.0581
Throughput Math	24.7(2.16)	30.8(2.29)	27.2(2.16)	F(2,23)=1.9, p=0.1717	

Sleep received in the previous 24 hours was negatively correlated with performance during the underway period at low sea states but not inport, a possible indication of poor quality sleep from fragmentation. Table 20 shows these correlations.

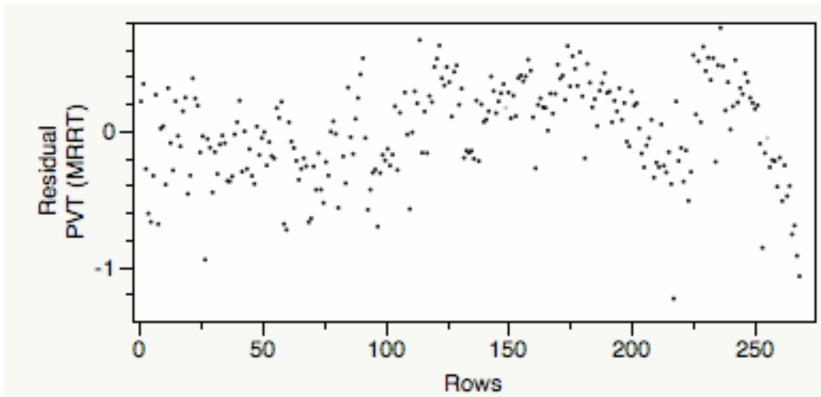
Table 20. Relationship between cognitive performance and sleep received in the

previous 24 hours, by sea state

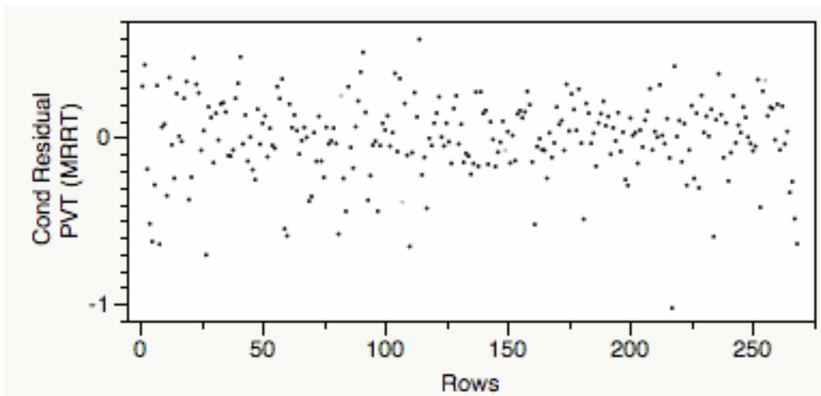
Performance Metric	Spearman Correlation
Inport	
MRRT Overall x Sleep	r_s=0.0549, p=0.5954
Throughput Overall x Sleep	r_s=0.0238, p=0.8181
Throughput Manikin x Sleep	r_s=0.0235, p=0.8205
Throughput Math x Sleep	r_s=0.0127, p=0.9024
Underway – Low Sea State	
MRRT Overall x Sleep	r_s=-0.4363, p=0.0027
Throughput Overall x Sleep	r_s=-0.4144, p=0.0046
Throughput Manikin x Sleep	r_s=-0.4137, p=0.0047
Throughput Math x Sleep	r_s=-0.4525, p=0.0018
Underway – High Sea State	
MRRT Overall x Sleep	r_s=-0.2067, p=0.1376
Throughput Overall x Sleep	r_s=-0.2223, p=0.1096
Throughput Manikin x Sleep	r_s=-0.2557, p=0.0646
Throughput Math x Sleep	r_s=-0.2142, p=0.1235

c. *Performance Regression Models*

Through a process of trial and error, various models were explored to test the effects of motion on performance at sea. Because of the significant within subjects variability among the participants (similar to that seen in the sleep models), a mixed effects model was used where individual differences were treated as a random effect and the other variables were treated as fixed effects, making the variation between participants more predictable. As an example, Figure 19 shows the improvement in the model fit when participants are treated as random effects rather than fixed effects.



a) Fixed model residual vs. row showing pattern



b) Mixed effects model residual vs. row showing no pattern

Figure 19. Comparison of residuals between the a) fixed model and the b) mixed-effects model

Many models were fit when assessing the performance data. These models included the same independent variables as the sleep models: sea state, shift type, rank, group type, and their interactions. Although many combinations proved significant, the performance models for Vigilance Performance and Cognitive Performance with the best fits based on the data collected are summarized in Tables 21 and 22.

Table 21. Vigilance performance random effects model summary with F-test statistic to test for effects

R^2= 0.62, R^2_{adj}=0.61	Coefficient Estimates (SE)	F-Ratio (df)	P-Value
Seas	Inport: 0.052 (.079) LSS: -0.050 (.078) HSS: -0.002 (.12)	F(2,248)=3.85	0.0227*
Rank	Jr. Enlisted: 0.039 (.13) Sr. Enlisted: -0.136 (.12) Officer: 0.097 (.13)	F(2,16)=1.09	0.360
Rank x Seas	Jr Enlisted x Inport: 0.095 (.14) Jr Enlisted x LSS: -0.046 (.14) Jr Enlisted x HSS: -0.050 (.14) Sr Enlisted x Inport: -0.095 (.12) Sr Enlisted x LSS: 0.038 (.12) Sr Enlisted x HSS: 0.056 (.13) Officer x Inport: -0.001 (.13) Officer x LSS: 0.007 (.13) Officer x HSS: -0.007 (.14)	F(4,248)=2.72	0.030*

(SE stands for Standard Error)

Table 22. Cognitive performance random effects model summary with F-test statistic to test for effects

a) Cognitive performance using MRRT for correct answers as response variable – data transformed using reciprocal method to stabilize variance ($y' = y^{-1}$)

R^2=0.91, R^2_{adj}=0.908	Coefficient Estimates (SE)	F-Ratio (df)	P-Value
Seas	Inport: 0.016 (0.024) LSS: -0.006 (0.024) HSS: -0.010 (0.024)	F(2,182)=19.264	<0.0001*
Rank	Jr. Enlisted: -0.008 (0.052) Sr. Enlisted: -0.055 (0.033) Officer: 0.062 (0.033)	F(2,9)=3.099	0.0948

b) Cognitive performance using overall throughput as response variable

R^2=0.931, R^2_{adj}=0.93	Coefficient Estimates (SE)	F-Ratio (df)	P-Value
Seas	Inport: -3.284 (3.6) LSS: 0.805 (3.6) HSS: 2.480 (3.6)	F(2,182) =9.069	<0.0001*
Shift Rotation	No Rotation: -6.569 (4.6) Yes Rotation: 6.569 (5.4)	F(1,10)=3.348	0.0970

These models show that both the vigilance and cognitive performance were reduced when underway. However, cognitive performance also showed increased throughput during higher sea states, again possibly due to the continued learning effect seen in the Switching test.

2. Actual Performance Compared to Predicted Performance Model

Performance was predicted using the FAST^{TM} interface based on the SAFTE model. For every time that the PVT or Switching test was actually taken, the predicted performance was derived using the FAST^{TM}. The predictions were based on the participant's most recent sleep history as determined by actigraphy sleep data as explained previously in the data and methodology section. The performance predictions are calculated in a percentage. The following analysis compared the actual performance results in the original units to the predicted in percentage out of ideal best level (100%). Additional analysis was conducted based on the assumption that the SAFTE model correctly predicted performance during the inport period where no motion was felt. In the additional analysis, actual performance was normalized based on the average inport period as a baseline equivalency (percent predicted equals actual performance during inport). If the model accurately predicted performance in all conditions, then no difference would be seen during the underway periods where increased motion was observed. Table 23 and 24 show the correlations between the performance, as predicted by the model, and the actual performance for both the vigilance and cognitive tests. In both cases, the inport period showed the highest correlation while the correlations during underway periods showed little relationship, indicating that the model did not adequately account for the factors affecting performance in an at sea environment.

Table 23. Correlation between predicted and actual vigilance performance (PVT)

Predicted Performance	Actual Performance	Actual Performance (% -normalized)
Inport	r=0.1338, p=0.1822	r=0.8423, p= <0.0001*
LSS	r=0.0018, p=0.9857	r=0.4828, p= <0.0001*
HSS	r=-0.1115, p=0.3764	r=0.6621, p= <0.0001*

* refers to statistical significance at the 0.001 level (p<0.001)

Table 24. Correlation between predicted and actual cognitive performance (Switching MRRT overall)

Predicted Performance	Actual Performance	Actual Performance (% -normalized)
Inport	r=0.4885, p= <0.0001*	r=0.4413, p= <0.0001*
LSS	r=0.0411, p=0.7886	r=0.4675, p=0.0012
HSS	r=0.0883, p=0.5292	r=0.2358, p=0.0891

* refers to statistical significance at the 0.001 level (p<0.001)

V. DISCUSSION

A. EFFECTS OF MOTION ON SLEEP AND PERFORMANCE

As seen in the results, sleep quality was reduced during underway periods due to an increase in activity during sleep periods. This finding was expected due to increased motion in the sleep environment. However, many variables beyond motion could account for the poor sleep quality and they were not looked at in this study. These factors include environment, pharmaceutical agents during higher sea state (resulting in increased sleepiness, but reduced sleep quality), caffeine intake, etc.

Higher overall daily sleep quantities were also seen during higher sea states. This finding may be due to the poor sleep quality achieved during high sea states, requiring additional overall time in bed to compensate to reduce fatigue symptoms as reported in the surveys. As discussed in the background section, motion sickness medication can cause drowsiness and may have been the reason for increased sleep quantity during the higher sea state periods. In addition, mild seasickness can cause sopite syndrome, characterized by lassitude, drowsiness, lack of motivation, and a minor state of depression. These factors could account for the increased rate of sleep during the high sea states as well.

Overall performance decreased with ship motion, but stabilized at high sea state. This result could be due to the increase in sleep quantity, which offset the motion effects. Additionally, learning continued throughout the trials and played a direct role in the performance results for the Switching test. Other factors that were not controlled for in this operational study, such as caffeine or medication intake, also could have affected the performance results.

Although we saw reduced quality and increased quantity of sleep during higher sea states, this study was limited to observations from actual operational trials. Unfortunately, there was no real baseline data, limiting our knowledge of typical sleep quality, quantity, and performance during optimal sleeping conditions. The inport period was used as a notional baseline; however, the ship was never fully motionless. Therefore, the inport period did not represent an ideal sleep environment. Additionally, the inport period was between two underway operational periods and involved significant corrective maintenance for the participants who stayed onboard and continued the trial, so the sailors could have been chronically sleep deprived with already altered sleep cycles.

B. DIFFERENCES BETWEEN FAST™ PREDICTIONS AND ACTUAL PERFORMANCE

1. Inport Versus Underway Correlations

When actual performance increased, predicted performance either stayed the same or decreased, indicating that the SAFTE model underestimates performance. The model could be overly sensitive to sleep disturbances or poor sleep quality. The finding indicates that the SAFTE model is accounting for the decrease in sleep, but overestimates the effects of poor quality sleep, or disturbances in sleep due to motion.

The FAST™ interface can account for sleeping environment by manually adjusting sleep quality. Unfortunately, the adjustments simply decreased the predicted performance even further. For increased accuracy during use in extreme work-sleep environments, the interface needs to properly account for the sleep disturbances. Currently, the FAST™ interface counts disturbances larger than 40 counts per second as fully awake. Alternatively, the data can be smoothed as continuous sleep at intervals of 5 minutes to 15 minutes. The model needs to be adjusted to reduce the effect of these disturbances due to ship motion on sleep quality, rather than to eliminate them.

Another reason that the performance predictions were inaccurate at all periods could be due to the interface's requirement for continuous sleep data. Any excluded period or gaps in the data extracted from the Actiware software was counted in FAST™ as awake unless manually adjusted. In this study, the periods missing actigraphy were filled in with manually entered log data. This discrepancy created variances in the resulting performance predictions across subjects. Future studies should strongly encourage the participants to wear their actiwatches continuously to control for this factor.

It has been shown that the rate at which recuperation occurs during sleep varies continually as a function of extant sleep debt (Hursh et al., 2004). When sleep debt is relatively high, then replenishment rate is higher during the beginning of sleep (Harrison, 1996; Lumley et al., 1986). This shift could be happening with the sailors who are constantly sleep deprived. Although sleep quality was poor, naps and short periods of sleep could have provided higher rates of reservoir replenishment than accounted for by the SAFTE model. The SAFTE model was validated on total and partial daily sleep deprivation in laboratory conditions, and has never been tested with sleep disturbances and the rejuvenating effect of napping. However, we are unable to account for the typical sleep patterns of the crew in order to determine root cause since there was no real baseline in this study.

2. SAFTE Model Improvements

The original hypothesis of this study was that the SAFTE model was not adequately compensating for the energy expenditure during wake periods in at sea environments, therefore underestimating the required replenishment to the sleep reservoir and over-predicting individual performance levels. Unfortunately, due to the poor implementation of the experiment due to unexpected external operational requirements, the controls were not in place to determine the full effect of ship motion without the individual compensation by increased sleep quantity at high sea states. Without controls, it is difficult to apply any changes to the model or make a quantitative recommendation for model improvement.

However, I am able to conclude from the data that the original theory was refuted by the results. The SAFTE model and FAST™ interface are inadequately predicting performance probably due to inaccurate calculations of reservoir replenishment, but not in the way first proposed. The model is not properly accounting for the sleep disturbances occurring in the maritime environment. The model should be adjusted to reduce the effects of these disturbances due to ship motion on the sleep quality by modifying the reservoir depletion rate or the reservoir replenishment portion of the model for maritime applications. Specifically, the compensation that the model uses for sleep fragmentation may be decreasing the restorative effect of sleep received by assuming that Stage One sleep will be achieved after every motion-induced wakening (Hursh et al., 2004).

As previously discussed, the sleep reservoir may be replenished faster during the shorter sleep periods during the underway high motion states, thus confounding the effects of motion on performance. In addition, the sleep fragmentation is currently offset in the model by eliminating five minutes of sleep after every wakening, as defined by motion over 40 activity counts per minute. During high seas, motion disturbances are increased during sleep, reducing sleep quality. However, these increased activity counts may not put the individual into a fully awake state as assumed by the SAFTE model. Humans are adaptable, and over time will become accustomed to a poor sleep environment in order to survive. It could be that the activity counts to determine wake periods are set too low, or the amount of time after a wake period required for Stage One sleep is set too high. The O'Hanlon et al. study in 1977 showed that sleep stages are altered during maritime environments, causing the individual to have a severely shortened Stage One cycle. In addition, even moderate sleep deprivation can alter one's sleep architecture, decreasing the amount of time in Stage One sleep (Van Dongen et al., 2003). When a human is deprived of certain sleep stages, they will jump to the required sleep stage almost immediately upon rest. Van Dongen et al., showed that even during partial sleep deprivation (4 hour and 6 hour per day), time in all stages but slow wave sleep (SWS) is significantly reduced (2003). This finding could be true of the sailors as well, since most of them are chronically sleep deprived due to high operational tempo.

C. LIMITATIONS — PROTOCOL AND SURVEY FIELDING METHODS REVIEW

This study had many limiting factors that hindered finding conclusive results. The main limitation was the implementation of the study. As with any field study, many external factors cannot be controlled. Due to unexpected maintenance issues, the ship in this study stayed in port during the middle of its operating period. This change in schedule caused some unexpected changes in the design of the study. The inport period was considered the "baseline" for the study; however, it was squeezed between two high-tempo underway operational periods. Additionally, the participants did not consistently participate during the inport or off-duty period. This inconsistency caused gaps in the data. It also reduced our ability to consider the inport period a truly controlled baseline for the study, where sleep levels and performance should have been at their peak.

The administration of the study and testing devices also were limiting factors on participation. The study was administered by a third party. The research team was not part of the military and had not previously gained the trust of the leadership or crew, causing hesitation among the crew to participate fully. Participation levels were inconsistent across the testing days, as shown in the data section. The crew was inconsistent in following the testers' instructions on use of the actiwatch, as well as recognizing the importance of using the activity log and taking the performance tests on a regular basis. All of these factors led to large variation between participants in the results. In addition, many of the aciwatch devices failed during the trials, reducing the number of participants even further.

Another problem was using the standardized methodologies and protocols when using the Actiware software and FAST™ interface. As discussed in the methodology section, the manner in which the sleep data is imported into FAST™ greatly effects the resulting predictions by the model. There was over 20% variance in the performance outcomes depending on the importing methodologies. This is a major concern for the validity and replication of previous and future studies using the FAST™ interface. It is additionally an issue if this program is used for establishing manpower and watchstanding requirements in operational settings. This does not even account for the additional variation resulting from the various importing and cleaning methods for the sleep data in the Actiware software program. Official recommended protocols for scoring actigraphy and exporting to FAST™ should be established in order to increase accuracy and prevent erroneous conclusions due to errors in effectiveness prediction.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSIONS AND RECOMMENDATIONS

A. EFFECTS OF MOTION ON SLEEP QUALITY

Based on the results of this study, it is apparent that the motion onboard a maritime platform has a negative effect on sleep quality. While at sea, activity counts during sleep periods increased significantly, indicating sleep fragmentation. This study population may have been chronically sleep deprived due to a high operational routine, which could negatively affect the crew's sleep-wake patterns as previously discussed. There is little that can be done to improve the sleep quality besides improving habitability through a ship hull and berthing compartment sleeping redesign to reduce the motion effects felt by passengers.

Future studies should account for the sleeping position relative to seas and ship motion. In addition, sleeping environmental factors such as noise, temperature, and humidity should be accounted for and controlled during future studies. Additionally, pharmaceutical, caffeine, and tobacco use should be controlled in future studies. Finally, future studies should collect baseline data in ideal sleeping conditions for comparison to ensure the participants are not suffering from total or partial chronic sleep deprivation, which may change their current sleep habits.

B. EFFECTS OF MOTION ON PERFORMANCE

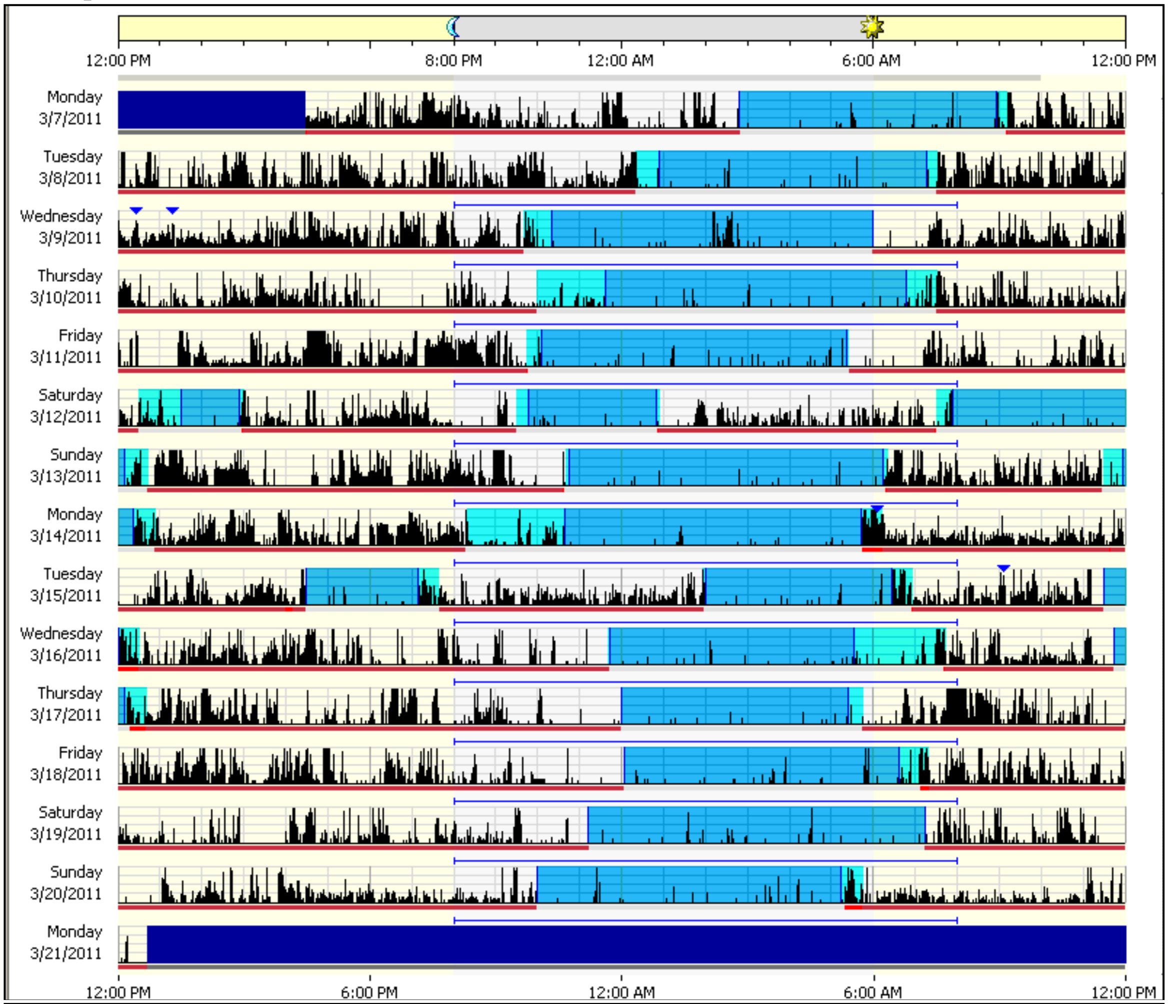
Performance deteriorated when participants were tested under motion. However, due to confounding variables and inadequate controls, it is unclear that there was a direct relationship between performance and motion. Other factors such as increased sleep duration during high sea states, use of seasickness medication, and caffeine use may have affected the performance results. In addition, there was no baseline data collected in an ideal, stationary environment for comparison purposes. Further studies in this area must use closely controlled experimental designs for conclusive results.

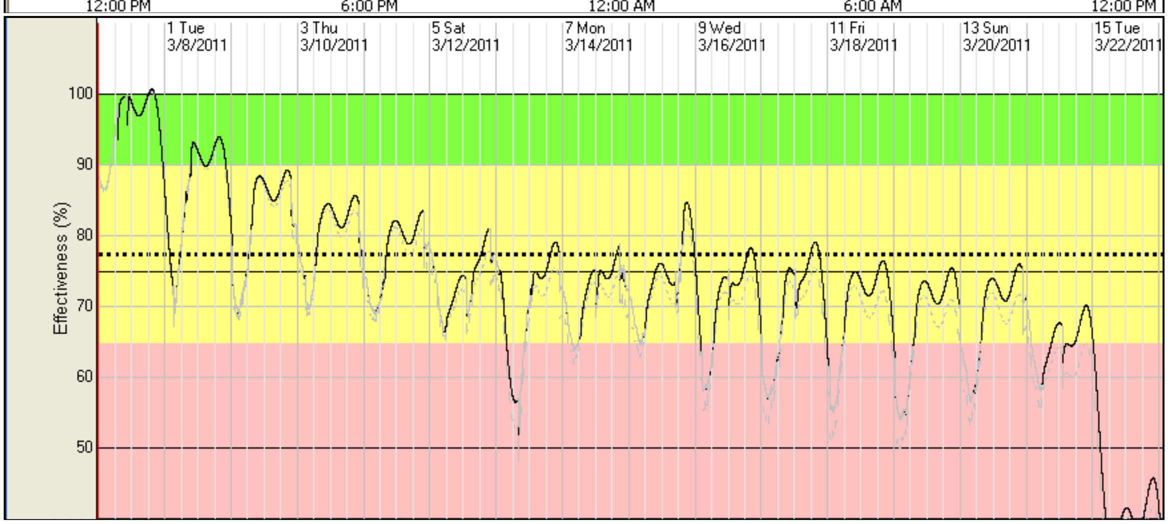
C. ACCURACY OF THE SAFTE MODEL FOR MARITIME ENVIRONMENT APPLICATIONS

The SAFTE model did not adequately predict performance in this study. The predictions were too low, overestimating the effects of poor sleep quality occurring during the underway period on performance. Recommendations for SAFTE model improvement discussed previously included a focus on changing the sleep reservoir replenishment rate by reducing the effect of sleep fragmentation during at sea operations. However, future studies need to explore this theory further in order to extrapolate the model findings to the entire maritime domain.

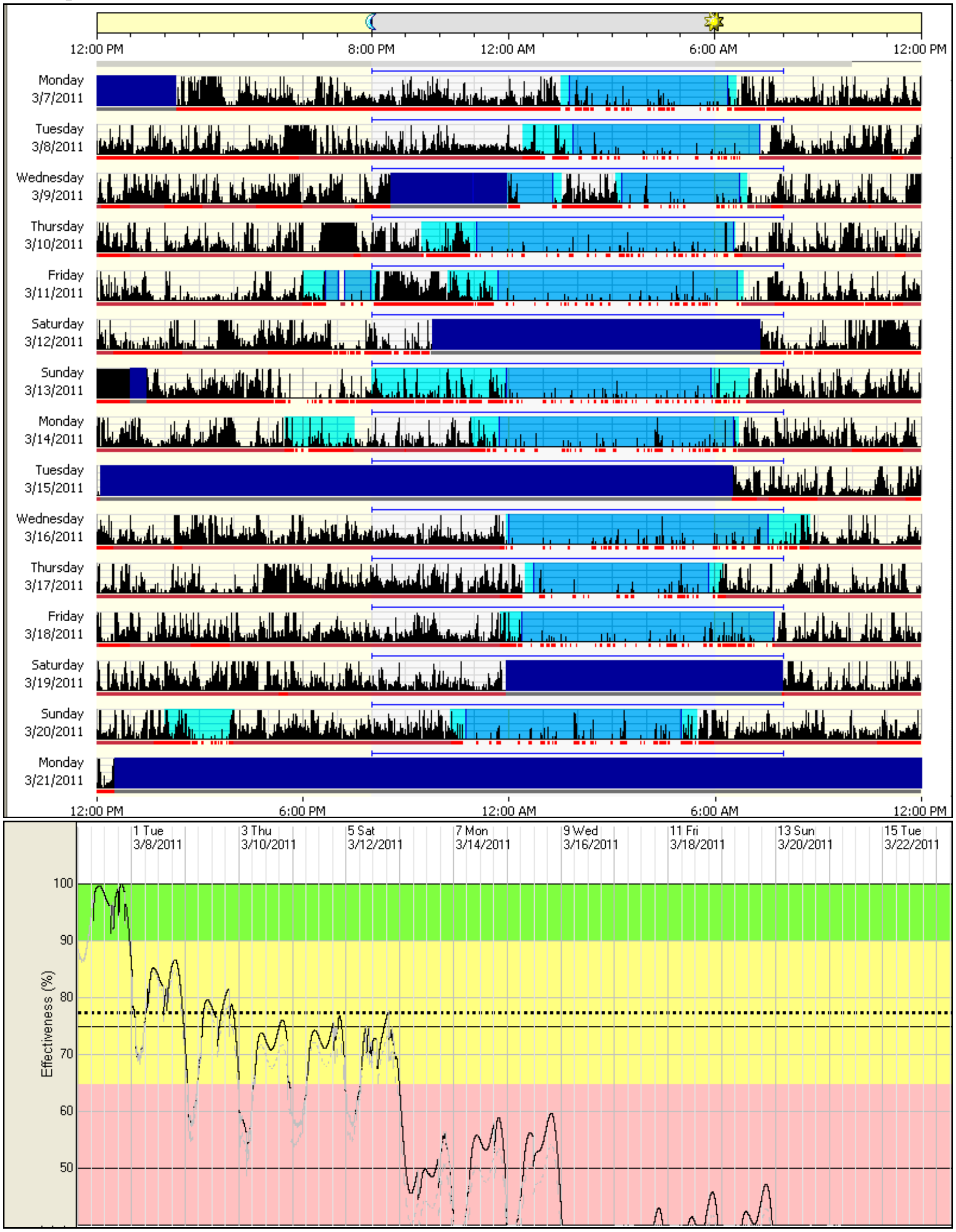
APPENDIX. ACTIGRAPHY DATA AND FAST™ ANALYSIS

Participant B572

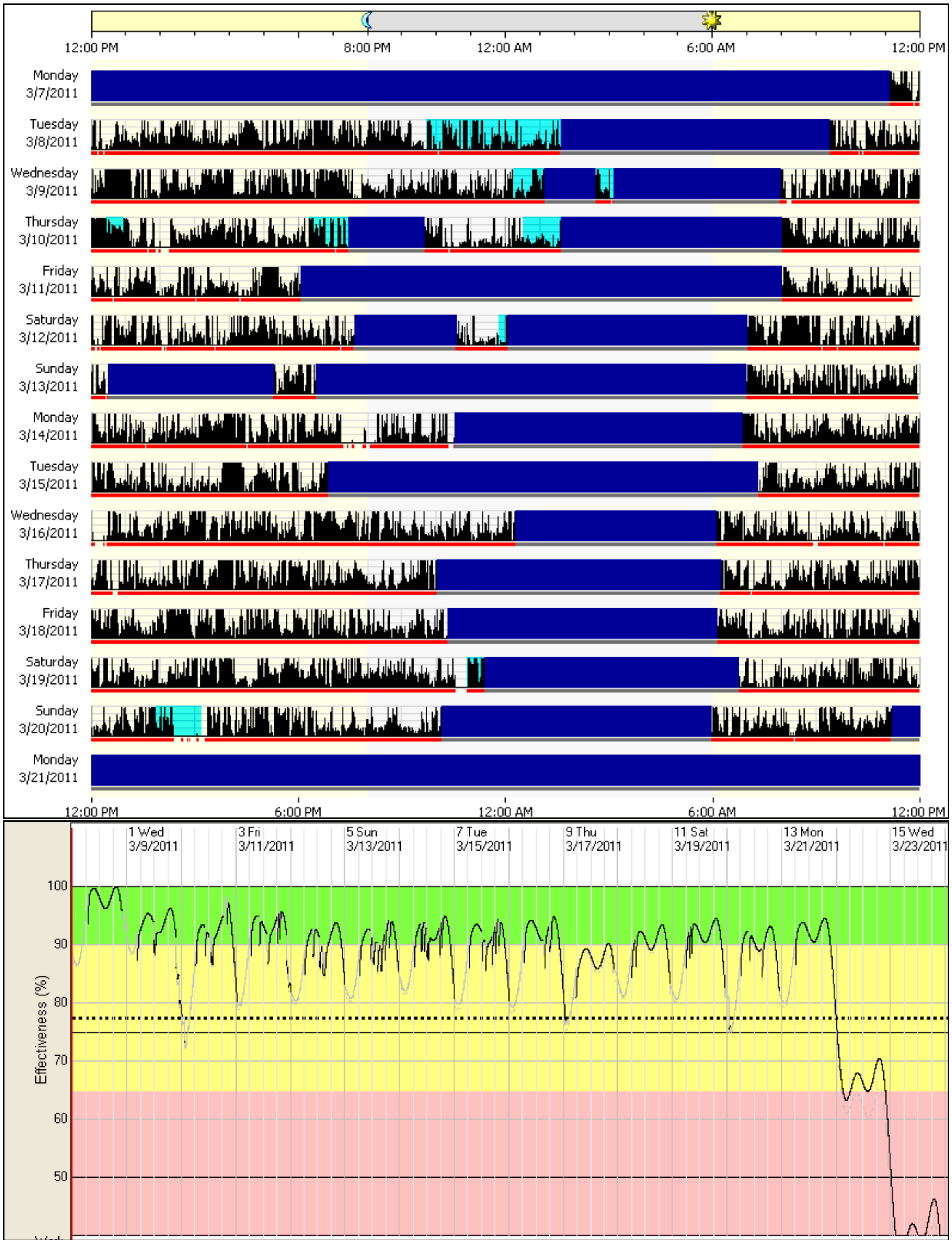




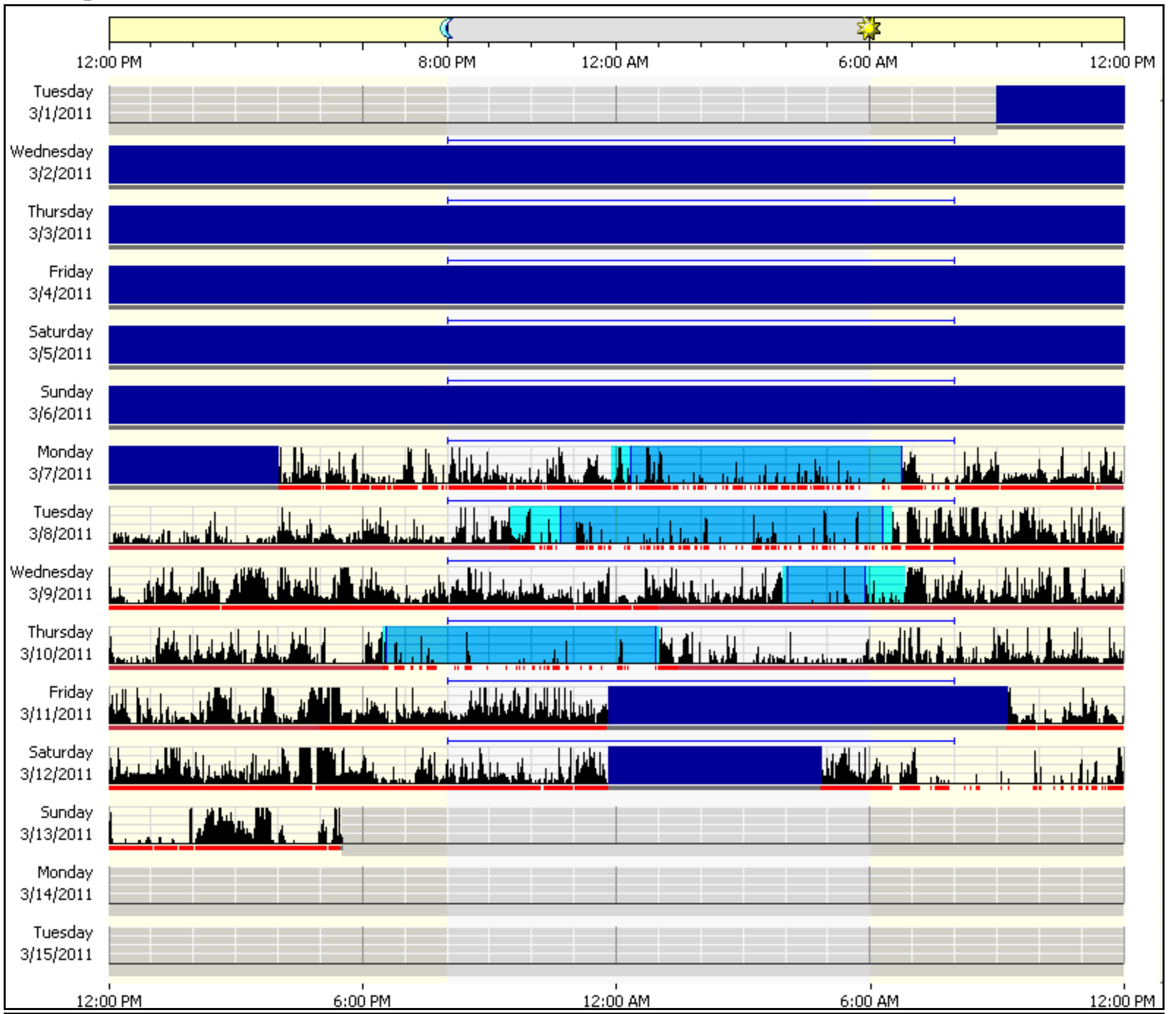
Participant C169

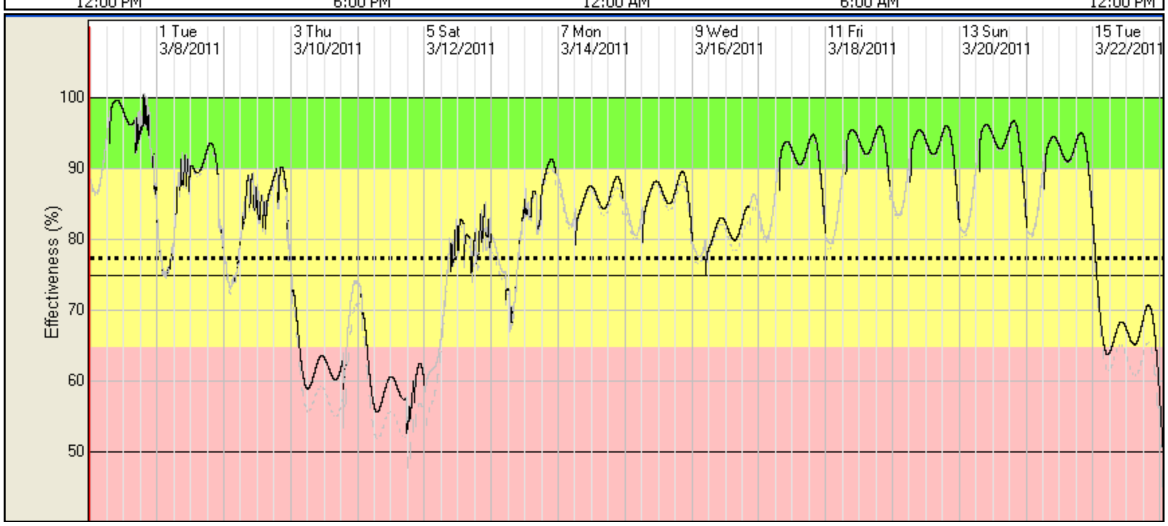


Participant C845

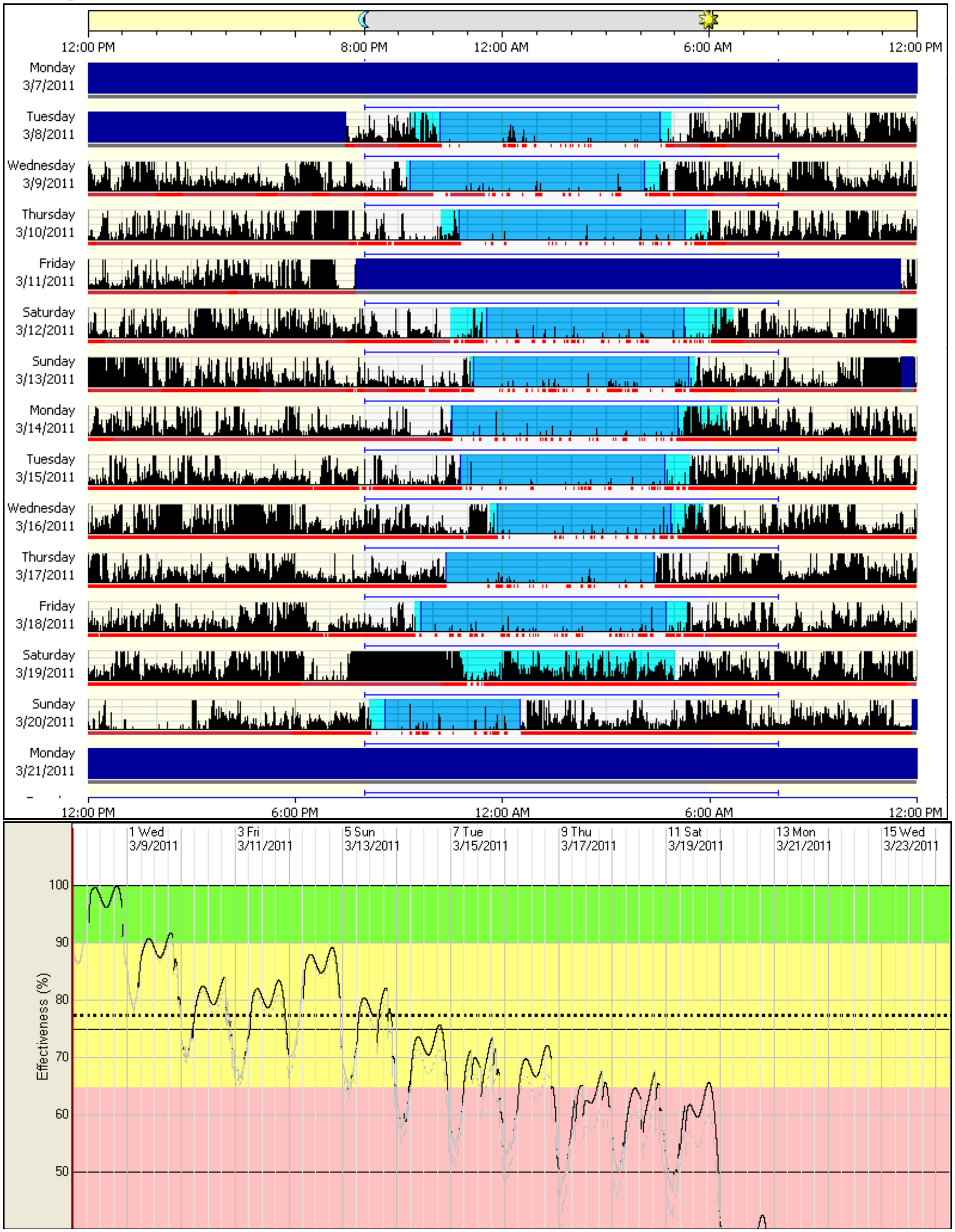


Participant G700

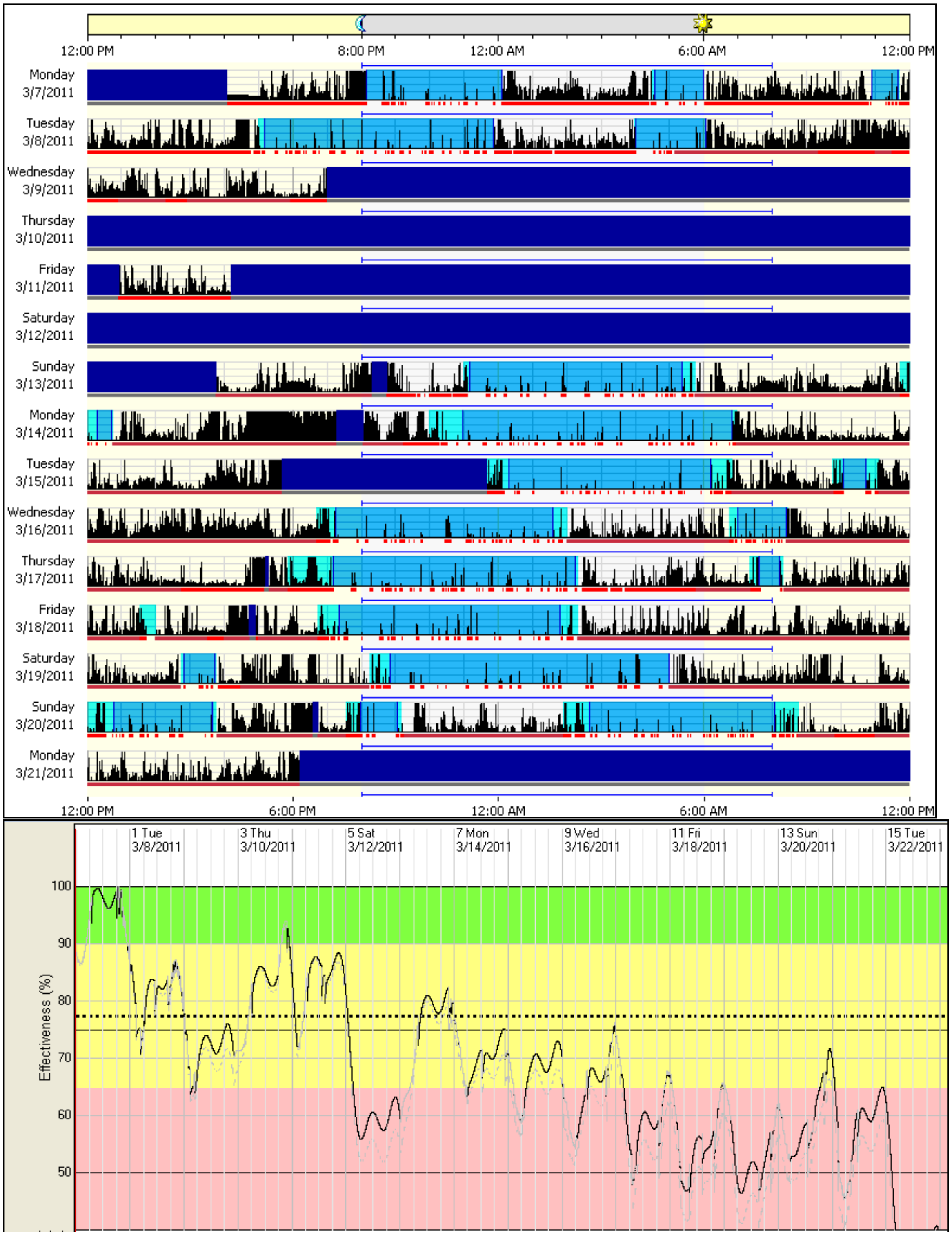




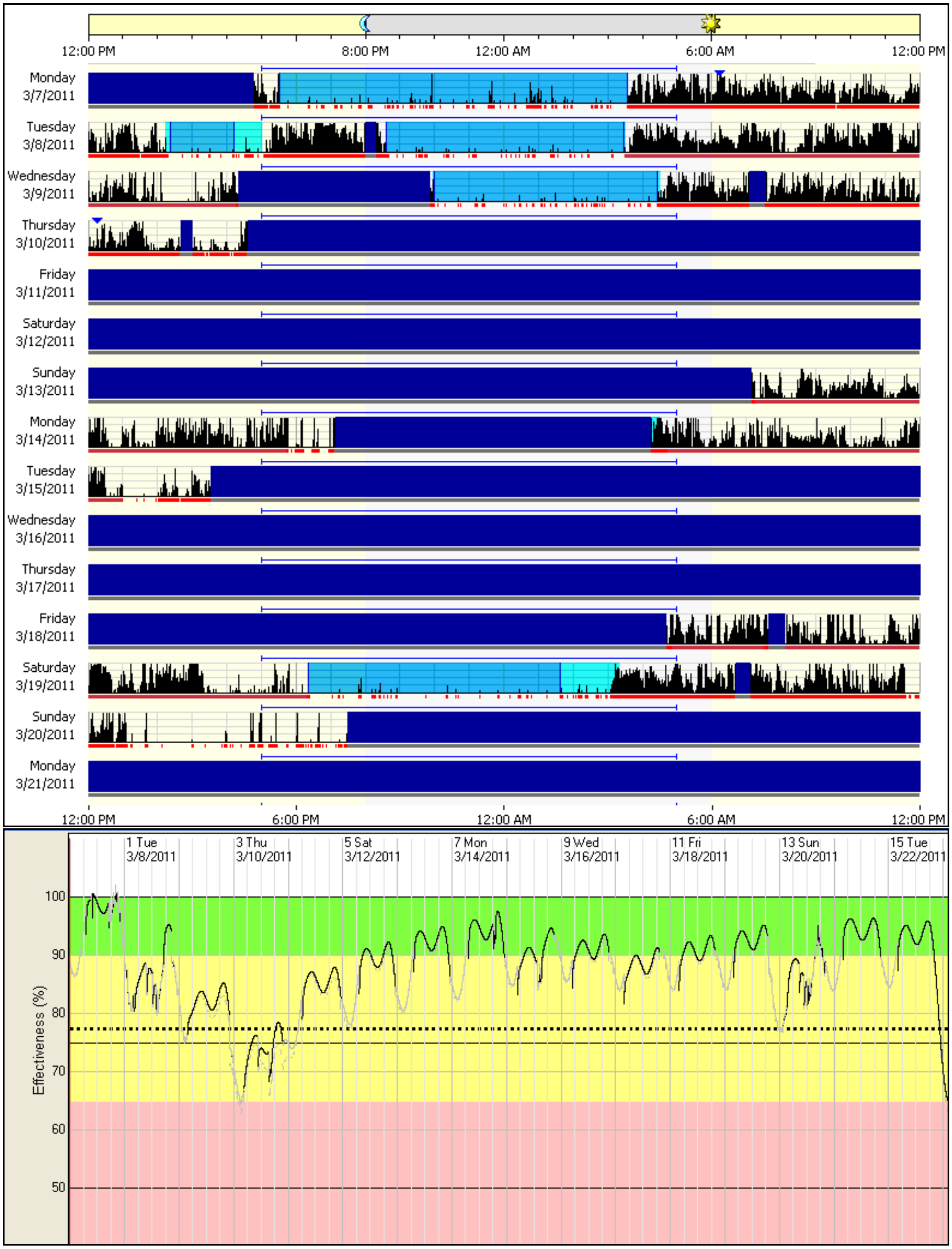
Participant I323



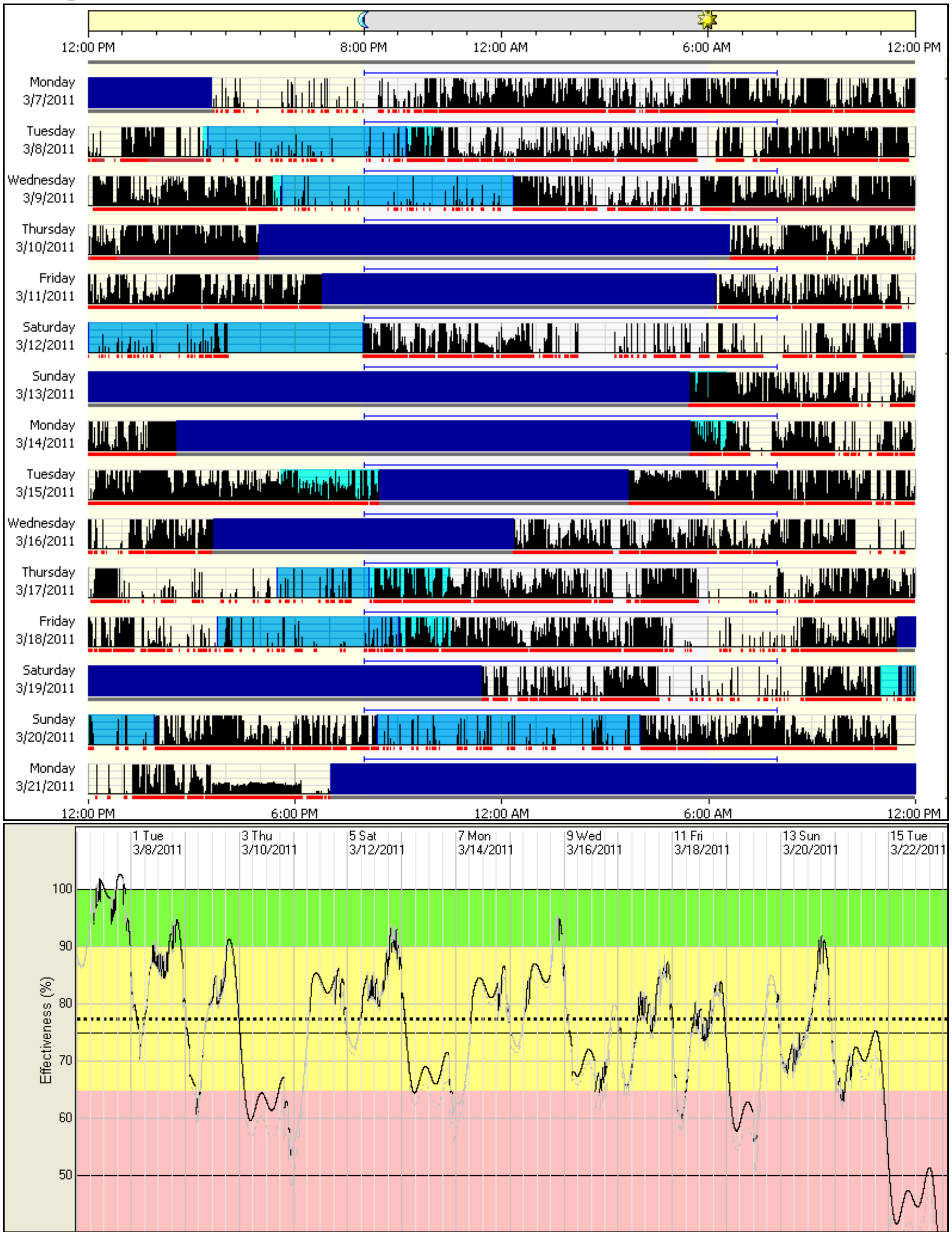
Participant I499



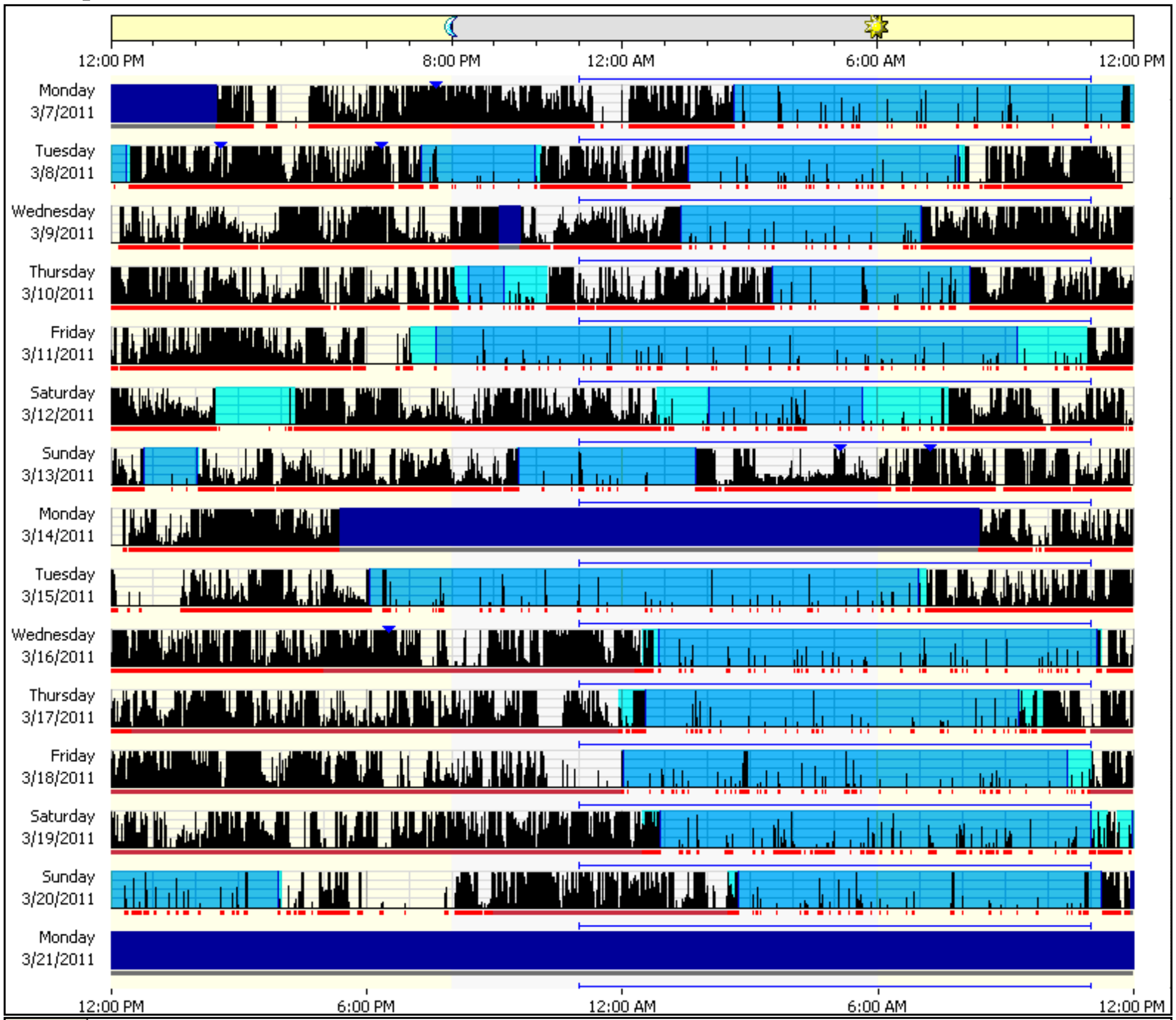
Participant K566

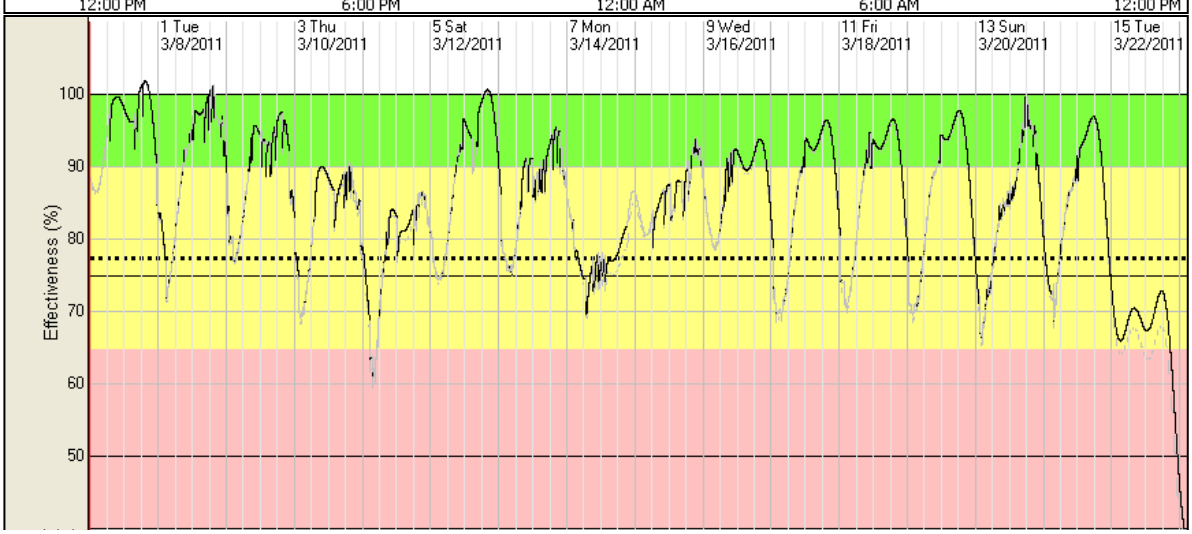


Participant K597

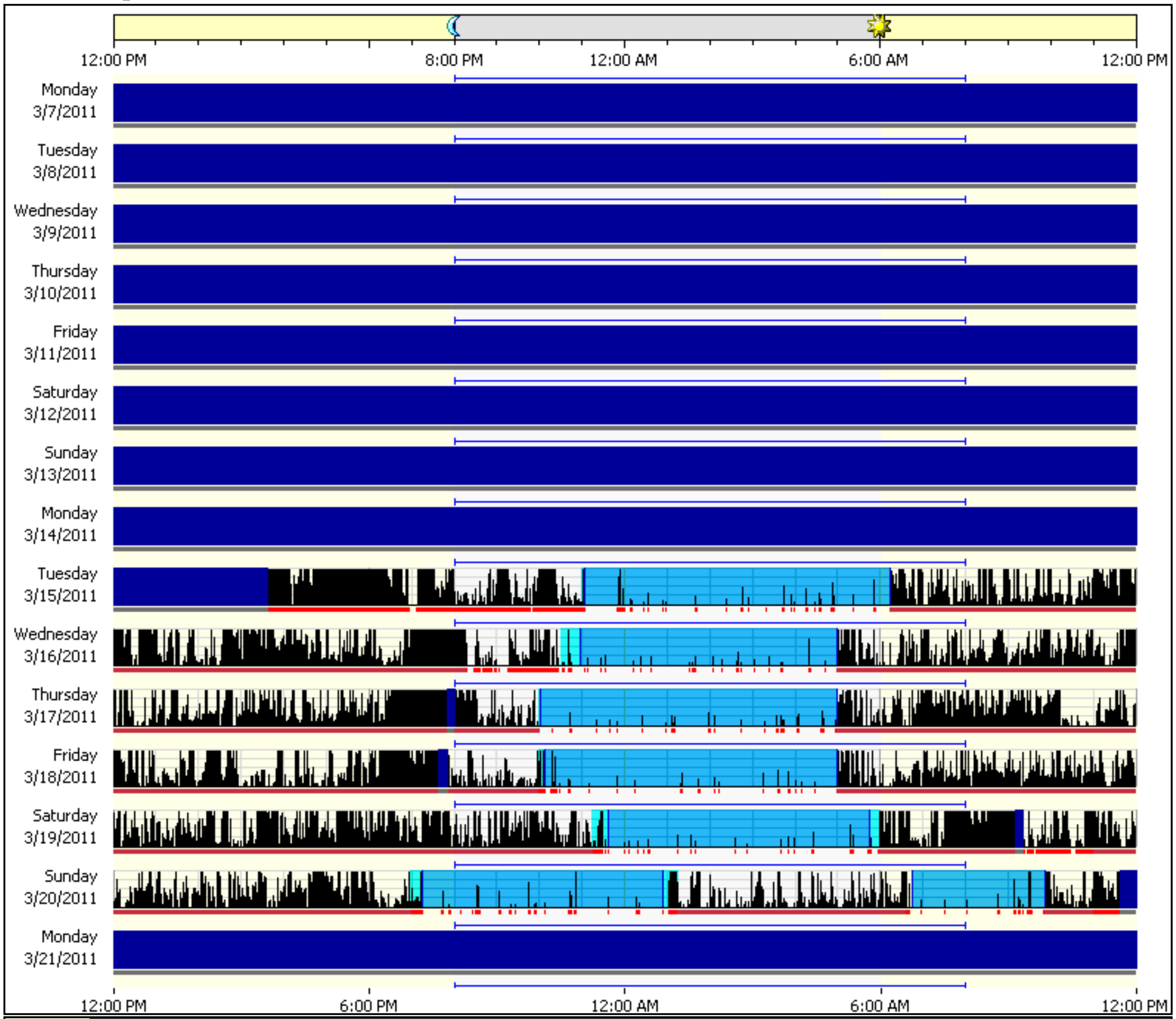


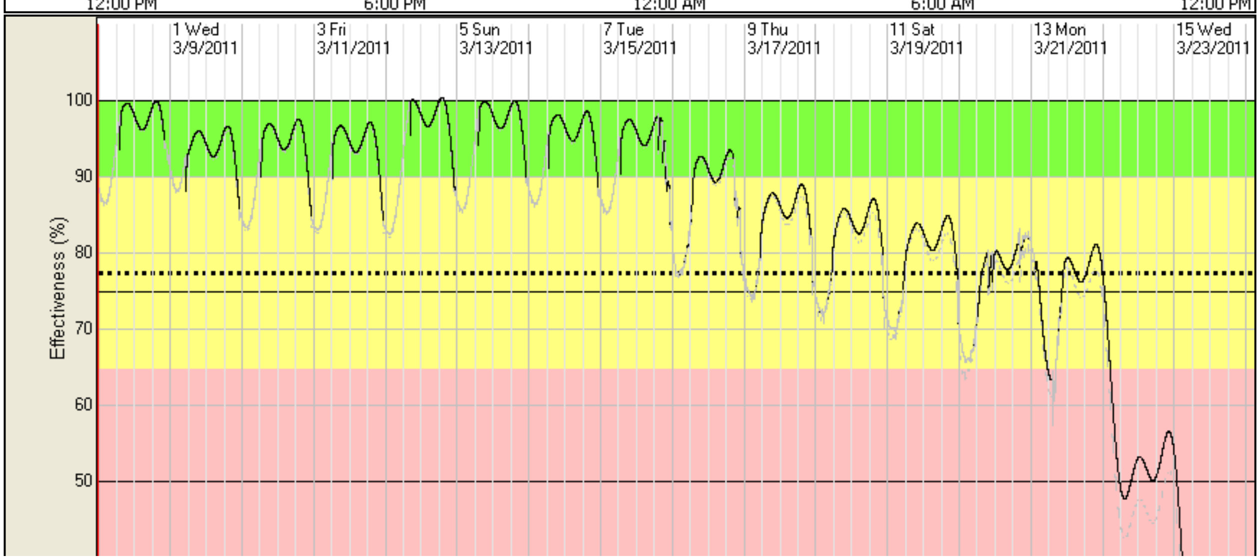
Participant K823



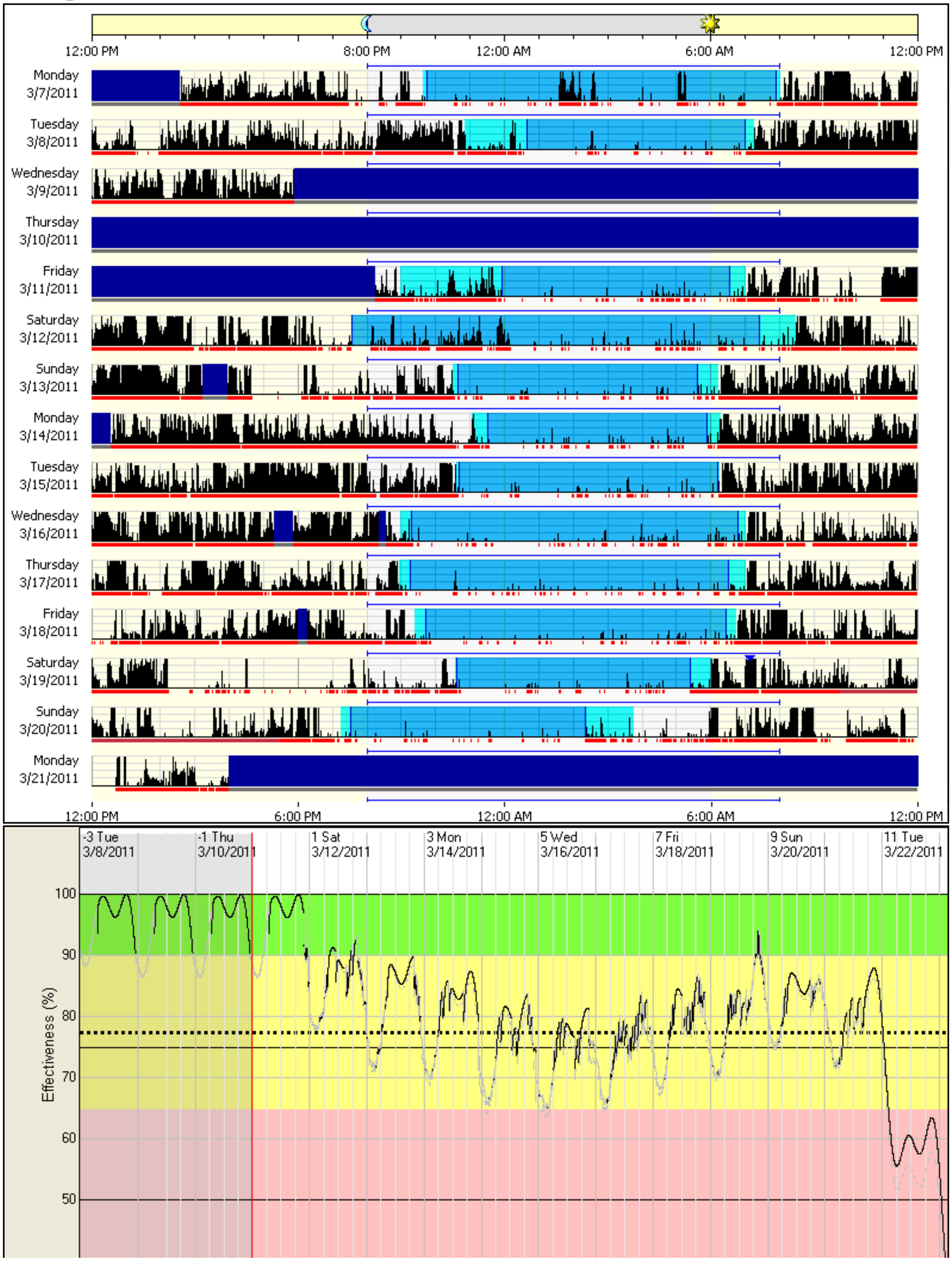


Participant N446

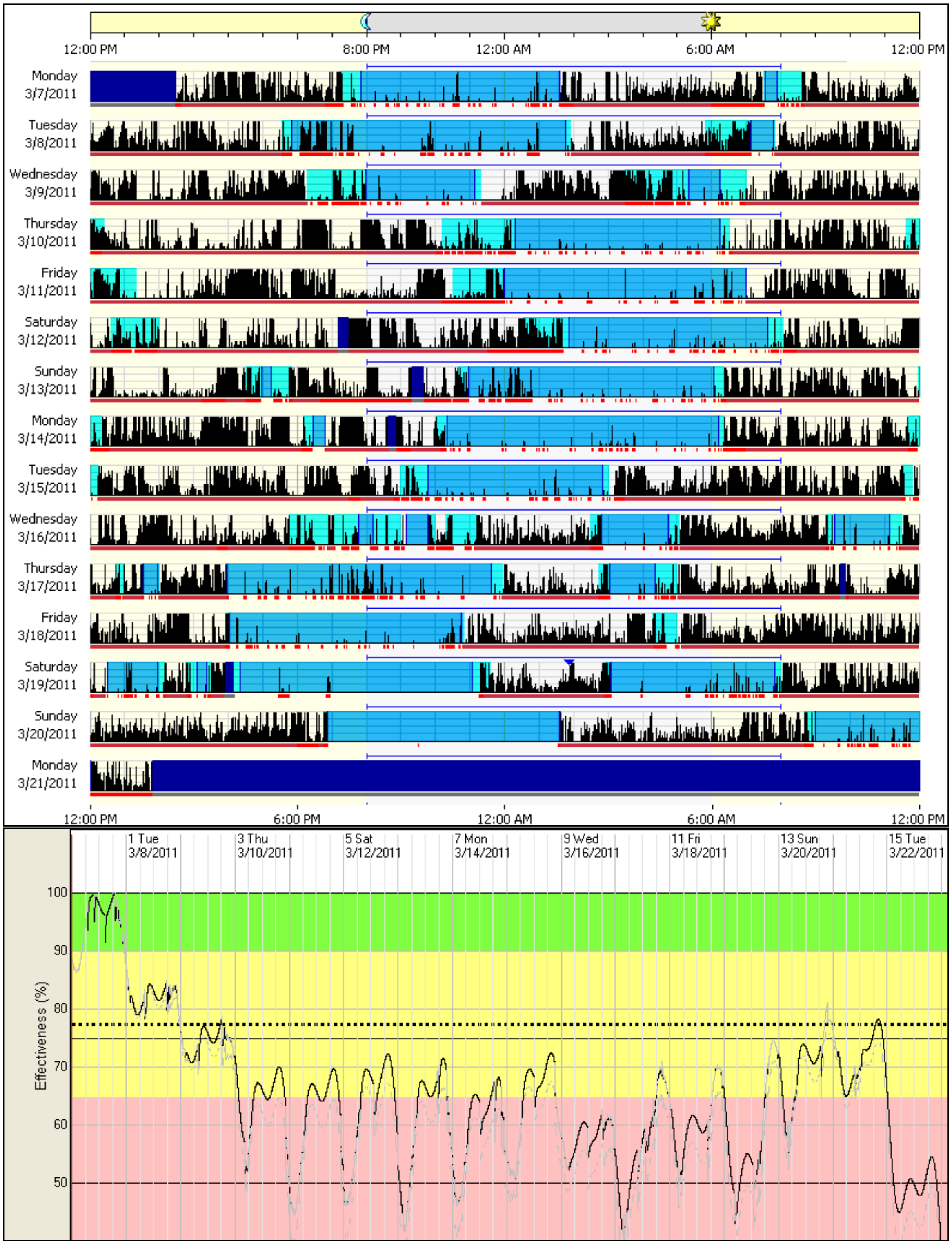




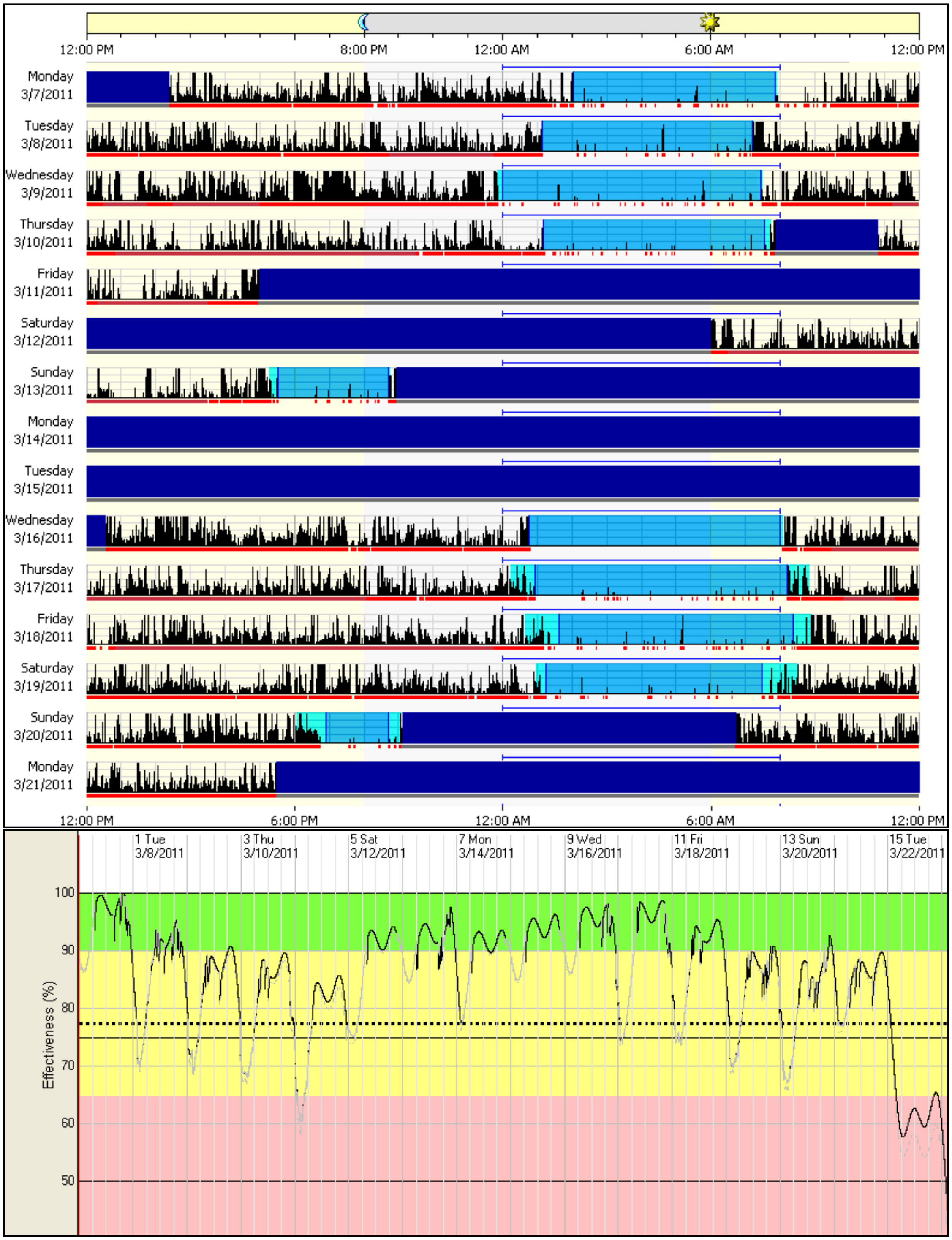
ParticipantR572



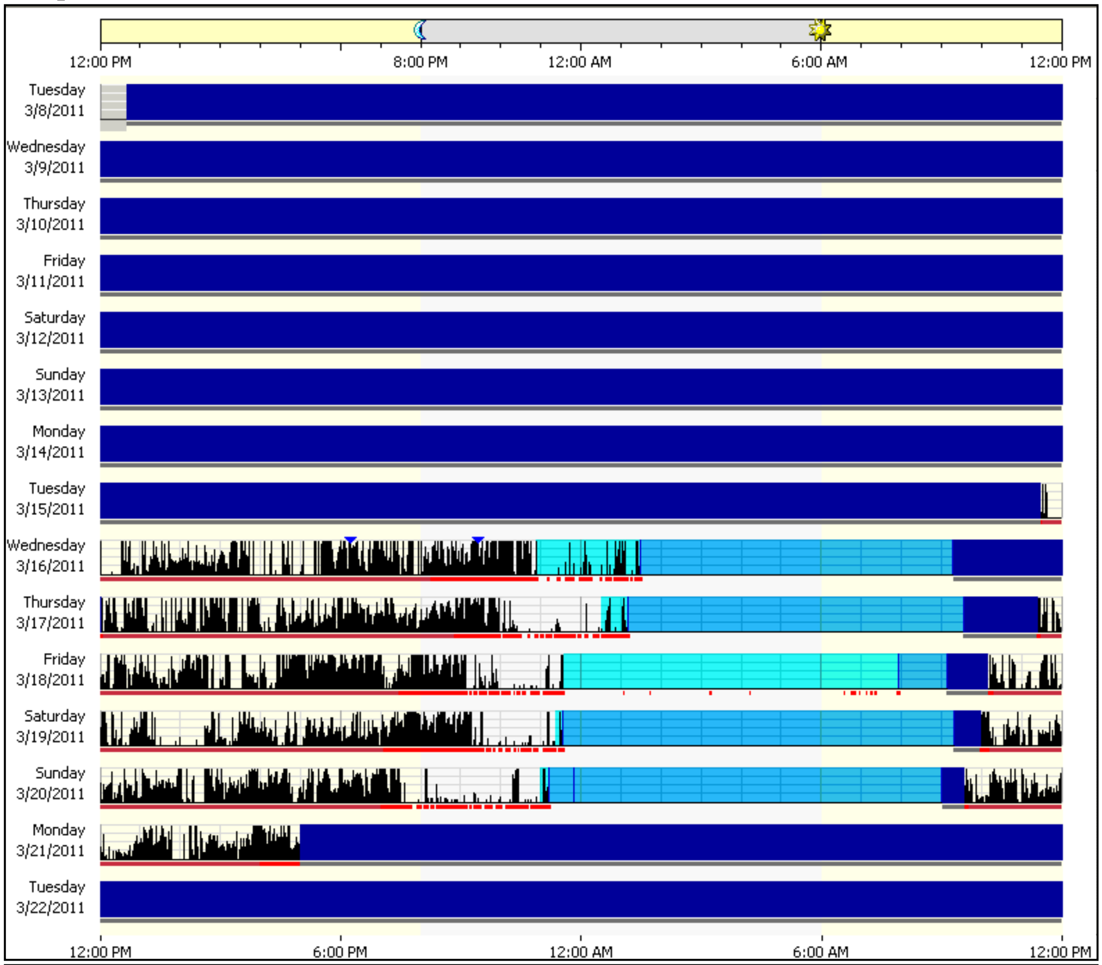
Participant T313

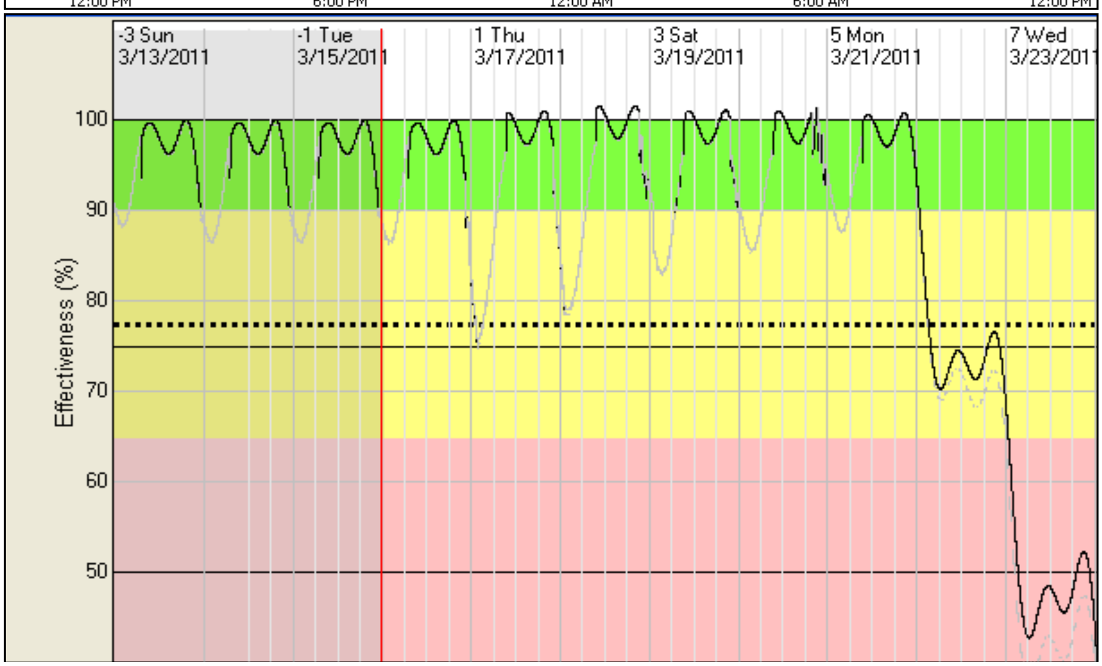


Participant V516

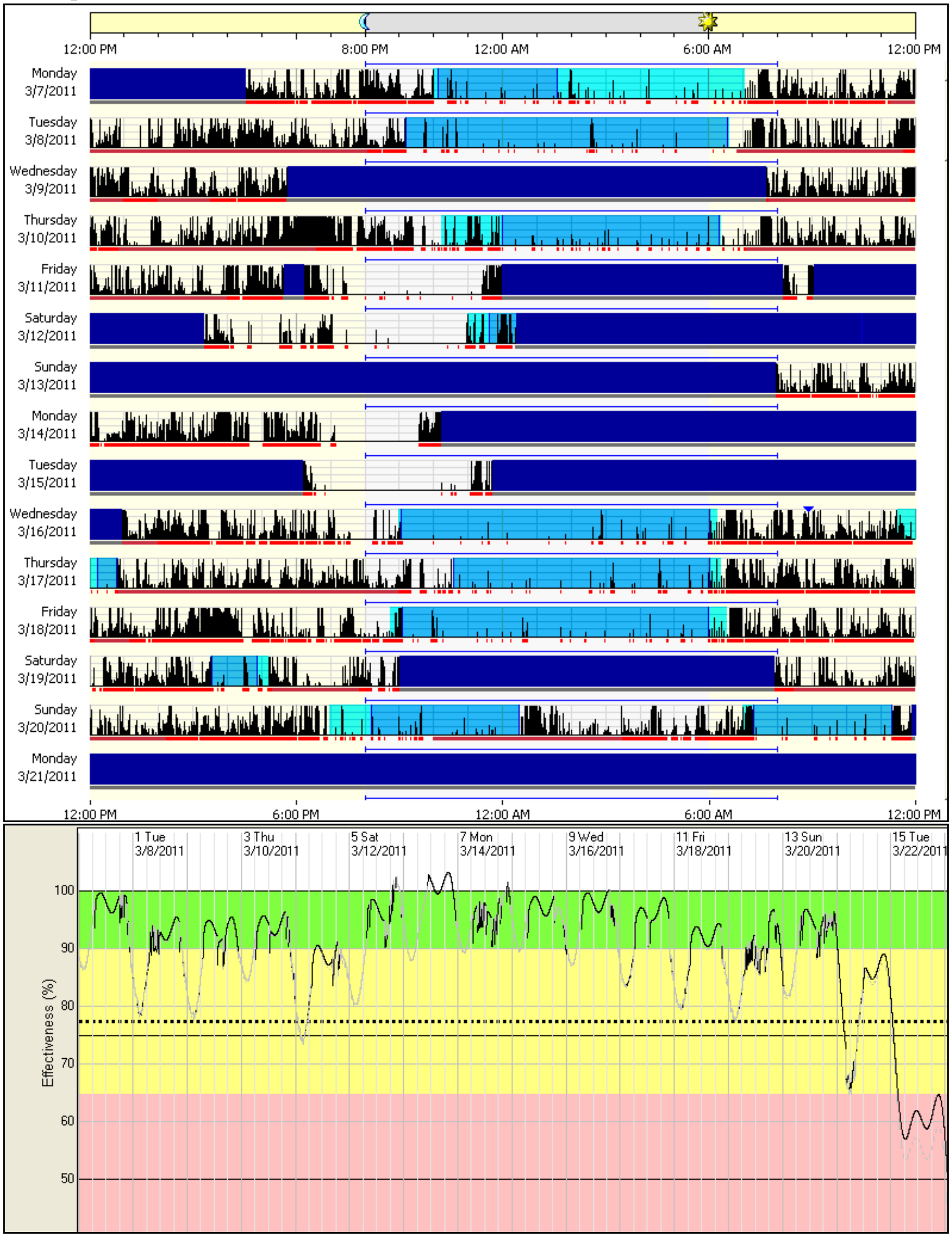


Participant X866





Participant Z772



THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Achermann, P., & Borbély, A.A. (2003). Mathematical models of sleep regulation. *Frontiers in Bioscience, 8,* 683–693. Retrieved from http://www.bioscience.org/u37153137/gaDTRQo7632rgysaGWQYT64356/2003/v8/s/1064/1064.pdf

Akerstedt, T., & Folkard S. (1997). The three-process model of alertness and its extension to performance, sleep latency, and sleep length. *Chronobiology International*, *14*(2), 115–123.

Akerstedt, T. (1990). Psychological and psychophysiological effect of shift work. *Scandinavian Journal for Work, Environment & Health*, *16*, 67–73.

Balkin, T., Thorne, D., Sing, H., Thomas, M., Redmond, D., Wesensten, N., Williams, J., Hall, S. & Belenky, G. (2000). *Effects of sleep schedules on commercial motor vehicle driver performance* (DOT-MC-00–133). Washington, DC: Department of Transportation, Federal Motor Carrier Safety Administration.

Basner, M., & Dinges, D. F. (2011). Maximizing Sensitivity of the Psychomotor Vigilance Test (PVT) to Sleep Loss. *Sleep, 34*(5), 581–591.

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., Russo, M. B., & Balkin, T. J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *Journal of Sleep Research, 12*, 1–12.

Brannen, Kate. (2012, January 26). Panetta: Military will be smaller but 'cutting edge.' *Defense News*. Retreived from http://www.defensenews.com/article/20120126/DEFREG02/301260014/Panetta-Military-Will-Smaller-8216-Cutting-Edge-8217-

Calhoun, S. R. (2006). *Human Factors in Ship Design: Preventing and Reducing Shipboard Operator Fatigue*. U.S. Costguard research project, Ann Arbor: University of Michigan. Retrieved from http://www.ardujenski.com/files/documents/FatigueDesign.pdf

Cole, R. J., Kripke, D. F., Gruen, W., Mullaney, D. J., & Gillin, J. C. (1992). Automatic sleep/wake identification from wrist activity. *Sleep, 15*(5), 461–469.

Colquhoun, P. W., & Folkard, S. (1985). Scheduling Watches at Sea. In Folkard, S., & Monk, T.H. (Eds.), *Hours of Work: Temporal Factors in Work-Scheduling*, 253–260. New York: John Wiley and Sons.

Coren, S. (1996). *Sleep thieves: An eye-opening exploration into the science and mysteries of sleep* (1^{st} ed.). New York: Simon & Schuster.

Crossland, P., & Lloyd, A. R. J. M. (1993). *Experiments to quantify the effects of ship motions on crew task performance: phase 1, motion induced interruptions an dmotion induced fatigue* (DRA/AWMH/TR/93025). Farnborough, UK: Defense Research Agency.

Defense Industry Daily (2012, May 23). LPD-17 San Antonio Class: The USA's New Amphibious Ships. *Defense Industry Daily*. Retreived from http://www.defenseindustrydaily.com/lpd17-san-antonio-class-the-usas-new-amphibious-ships-updated-02322/

Eddy, D. R., & Hursh, S. R. (2006). *Fatigue avoidance scheduling tool (FAST) phase II SBIR final report, part 1* (AFRL-HE-BR-TR-2006–0015). Dayton, OH: Air Force Research Lab, Human Effectiveness Division.

Eddy, D. R., Schiflett, S. G., Schlegel, R. E., & Shehab, R. L. (1998). Cognitive performance aboard the life and microgravity spacelab. *Acta Astronautica*, *43*, 193–210.

Folkard, S., & Monk, T. H. (1985). Circadian performance rhythms. In Folkard, S., & Monk, T.H. (Eds.), *Hours of Work: Temporal Factors in Work-Scheduling*, 37–52. New York: John Wiley and Sons.

Gillard, A. W. K., & Wientjes, J. E. (1994). Mental workload and stress as two types of energy mobilization. *Work and Stress, 8*, 141–152.

Goh, V. H., Tong, T. Y., Lim, C., Low, E. C., & Lee, L. K. (2000). Circadian Disturbances after night-shift work onboard a naval ship. *Military Medicine, 165* (2), 101–105.

Graw, P., Werth, E., Krauchi, K., Gutzwiller, F., Cajochen, C., & Wirz-Justice, A. (2001). Early Morning Melatonin Administration Impairs Pychomotor Vigilance. *Behavioural Brain Research, 121*, 167–172.

Hakola, T., & Härmä, M. (2001). Evaluation of a fast forward rotating shift schedule in the steel industry with a special focus on ageing and sleep. *Journal of Human Ergology*, *30*(1–2), 315–319.

Harvile, D., Harrison, R., & Scott, C. (2007). *The fatigue equivalent of job experience and performance in sustained operations* (AFRL-HE-BR-TR-2006–0042). Brooks City-Base, TX: Air Force Research Lab, Human Effectiveness Directorate, Biosciences and Protection Division, Biobehavioral Performance Branch.

Heldsdingen, A. S. (1996). *Sea movement and task performance*. The Netherlands: Free University of Amsterdam, Department of Psychology.

Heus, R., Wertheim, A., & Havenith, G. (1998). Human Energy Expenditure when walking on a moving platform. *European Journal of Applied Physiology: Occupational Physiology, 77* (4), 388–394.

Horne, J. A. (1985). Sleep Loss: Underlying mechanisms and tiredness. In Folkard, S., & Monk, T.H. (Eds.), *Hours of Work: Temporal Factors in Work-Scheduling*, 53–64. New York: John Wiley and Sons.

Hursh, S. R. (2003, September 11). *The Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE™) Model and Fatigue Avoidance Scheduling Tool (FAST™): Fatigue and alertness management using the SAFTE™ model and FAST™* [PowerPoint]. Retrieved from http://www.nps.navy.mil/orfacpag/resumePages/projects/Fatigue/HurshSAFTEFAST.pdf

Hursh, S. R., Redmond, D. P., Johnson, M. L., Thorne, D. R., Belenky, G., Balkin, T. J., Storm, W.F., Miller, J.C., & Eddy, D.R. (2004). Fatigue Models for Applied Research in Warfighting. *Aviation, Space, and Environmental Medicine, 75* (3Suppl.), A44-A53.

Jean-Lousie, G., Kripke, D.F, Cole, R.J., Assmus, J.D., & Langer, R.D. (2001). Sleep dectectino with an accelerometer actigraph: comparisons with polysomnography. *Physiology and Behavior, 72*(1), 21–28.

Johnson, M. L., Belenky, G., Redmond, D. P., Thorne, D. R., Williams, J. D., Hursh, S. R., & Balkin, T. J. (2004). Modulating the homeostatic process to predict performance during chronic sleep restriction. *Aviation, Space, and Environmental Medicine, 00*93, Suppl.), 1–6.

Kripke, D. F., Hahn, E. K., Grizas, A. P., Wadiak, K. H., Loving, R. T., Poceta, J.S., Shadan, F.F., Cronin, J.W., & Kline, L.E. (2010). Wrist actigraphic scoring for sleep laboratory patients: algorithm development. *Journal of Sleep Research*, *19*, 612–619. doi: 10.1111/j.1365–2869.2010.00835.x

Lieberman, H. R., Tharion, W. J., Shukitt-Hale, B., Speckman, K. L., & Tulley, R. (2002). Effects of caffeine, sleep loss, and stress on cognitive perofrmance and mood during U.S. Navy SEAL training. Sea-Air-Land. *Psychopharmacology*, *164*(3), 250–61.

Lumley, M., Roehrs, T., Zorick, F., Lamphere, J., & Roth, T. (1986). The alerting effects of naps in sleep-deprived subjects. *Psychophysiology, 23*(4), 403–8.

Luna, T. D., French, J., & Mitcha, J. L. (1997). A study of USAF air traffic controller shiftwork: sleep, fatigue, activity, and mood analyses. *Aviation, Space, and Environmental Medicine*, *68*(1), 18–23.

Mallis, M. M., Mejdal, S., Nguyen, T. T., & Dinges, D. F. (2004). Summary of the key features of seven biomathematical models of human fatigue and performance. *Aviation, Space, and Environmental Medicine, 75*(3 Suppl.), A4-A14.

Meere, E. P., & Grieco, L. (1996). Ship habitability – Preparing for the 21st century. *Association of Scientists & Engineers of the Naval Sea System Command: Annual Technical Symposium Technical Papers*, *33*, 132–139.

Naitoh, P., Kelly, T. L., & Englund, C. (1990). Health effects of sleep deprivation. *Occupational Medicine: State of the Art Reviews*, *5*(2), 209–237.

O'Hanlon, J. F., Miller, J. C., & Royal, J. W. (1977). *Effects of simulated surface effect ship motions on crew habilitability – Phase II, volume 4. Crew cognitive functions, physiological stress, and sleep* (PMS-304 TR 1070). Bethesda, MD: Naval Sea Systems Command.

Purnell, M. T., Feyer, A. M., & Herbison, G. P. (2002). The impact of a nap opportunity during the night shift on the performance and alertness 12-h shift workers. *Journal of Sleep Research*, *11*, 219–227.

Pulsar Informatics (2012). *PVT.* Retrieved from http://www.pulsarinformatics.com/pvt.html

Quant, J. R. (1992). The effect of sleep deprivation and sustained military operations on near visual performance. *Aviation, Space, and Environmental Medicine*, *63*(3), 172–176.

Reeves, D. L., Bleiberg, J., Roebuck-Spenser, T., Cernich, A.N., Schwab, K., Ivins, B., Salazar, A.M., Harvey, S.C., Brown, F.H., & Warden, D. (2006). Reference values for performance on the automated neuropsychological assessment metrics V3.0 in an active duty military sample. *Military Medicine*, *171*(10), 982–994.

Reeves, D. L., Winter, K.P., Bleiberg, J., & Kane, R.L. (2007). ANAM® genogram: Historical perspectives, desciptions, and current endeavors. *Archives of Clinical Neuropsychology*, *22S*, S15-S37.

Respironics, Inc. (2009). Actiware and actiware CT: Software and hardware manual. Retrieved from http://www.respironics.com/UserGuides/Actiwatch_SW_Instr_Manual.pdf

Rothblum, A. M. (2002). Human Error and Maritime Safety. *Second International Workshop on Human Factors in Offshore Operations: Working group 1, Human Factors in Incident Investigation and Analysis.* Retrieved from http://www.hsdl.org/?view&did=449631

Rutenfranz, J., Plett, R., Knauth, P., Condon, R., De Vol, D., Flethcher, N., Eickhoff, S., Schmidt, K.H., Donis, R., & Colquhoun, W.P. (1988). Work at sea: a study of sleep, and circadian rhythms in physiological and psychological functions, in watchkeepers on merchant vessels. II. Sleep. *International Archives of Occupational and Enivornmental Health*, *60*(5), 331–9.

Sack, R. L., Auckley, D., Auger, R. R., Carskadon, M. A., Wright, K. P., Vitiello, M. V., & Zhdanova, I. V. (2007). Circadian rhythm sleep disorders: Part I, basic principles, shift work and jet lag disorders. *Sleep, 30*(11), 1460–1483.

Sallinen, M., Härmä, M., Akerstedt, T., Rosa, R., & Lillqvist, O. (1998). Promoting alertness with a short nap during a night shift. *Journal of Sleep Research, 7*, 240–247.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors*, *37*(1), 5–19.

Schweitzer, P. K., Randazzo, A.C., Stone, K., Erman, M., & Walsh, J. K. (2006). Laboratory and field studies of naps and caffeine as practical countermeasures for sleep-wake problems associated with night work. *Sleep*, *29,* 39–50.

U.S. Library of Congress, Congressional Research Service. (2011, April 29). *Navy Littoral Combat Ship (LCS) Program: Background, Issues, and Options for Congress* by Ronald O'Rourke (CRS Report No. RL33741). Washington D.C.: Office of Congressional Information and Publishing.

Van Dongen, H. P. A., Maislin, G., Mullington, J. M, & Dinges, D. F. (2003). The cumulative cost of additional wakefulness: Dose-response effects on neurobehavior functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep 26*(2), 117–126.

Wertheim, A. H., & Kistemaker, J. A. (1997). *Task performance during simulated ship movements* (TM-97-A014). Soesterberg, The Netherlands: TNO Human Factors Research Institute.

Wertheim, A. H., Heus, R., & Vrijkotte, T. G. M. (1995). *Human energy expenditure, task performance, and seasickness during simulated ship movements* (TM-1995-C-29). Soesterberg, The Netherlands: TNO Human Factors Research Institute.

Wertheim, A. H. (1998). Working in a moving environment. *Ergonomics*, *41*(12), 1845–1858.

Woods, D. D., Sarter, N., & Billings, C. (1997). Automation surprises. In G. Salvendy (Ed.), *Handbook of Human Factors*, 1926–1943. New York: John Wiley.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Dr. Nita Lewis Shattuck
   Naval Postgraduate School
   Monterey, California

4. President Daniel T. Oliver
   Naval Postgraduate School
   Monterey, California